DOCUMENT RESUME

ED 051 288 TM 000 604

AUTHOR	Freeberg, Norman E.
TITLE	Development of Evaluation Measures for Use with Neighborhood Youth Corps Enrollees. Final Report.
INSTITUTION	Educational Testing Service, Princeton, N.J.
SPONS AGENCY	Department of Labor, Washington, D.C.
REPORT NO	PR-68-5
PUB DATE	Dec 68
NOTE	143p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$6.58
DESCRIPTORS	*Attitude Tests, Culturally Disadvantaged, *Disadvantaged Youth, Dropout Attitudes, Dropout Programs, Factor Analysis, Job Skills, Job Training, Minority Groups, *Nonverbal Tests, Occupational Aspiration, Oral Communication, *Program Effectiveness, Rating Scales, Sex Differences, Test Bias, Test Construction, Verbal Ability, Visual Measures, *Vocational Interests, Vocational Training Centers, Youth Opportunities
IDENTIFIERS	*Neighborhood Youth Corps

ABSTRACT

This study represents an initial phase in the development and application of a battery of measures for research and evaluation with disadvantaged adolescents in a youth-work training program. The paper-and-pencil instruments devised attempt to compensate for deficiencies in currently available tests by shaping content, format and administration to the academic level and cultural background of the intended respondents. A battery of 13 test booklets, incorporating such characteristics as relatively low verbal levels, oral presentation and, in many cases, pictorial formats, dealt with areas of vocational orientation (plans, aspirations, knowledge and interests); attitudes (toward self, authority and social values) and job-related reasoning skills. Based upon analyses of item characteristics, reliability estimates, validity (against proficiency ratings), factor analysis of the battery and research-related findings, a number of the measures were shown to possess a pattern of psychometric characteristics that warrants further research on their refinement and application. (Author)

PR-68-5

DEVELOPMENT OF EVALUATION MEASURES FOR USE WITH NEIGHBORHOOD YOUTH CORPS ENROLLEES

Norman E. Freeberg

U.S. DEPARTMENT OF HEALTH, EDUCATION
& WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

December 1968

EDUCATIONAL TESTING SERVICE
PRINCETON, NEW JERSEY

ED051288

Development of Evaluation Measures for Use with Neighborhood Youth Corps Enrollees

Norman E. Freeberg
Project Director
Educational Testing Service
Princeton, N. J.

Final Report
U. S. Dept. of Labor Contract #66-00-09
December, 1968

Acknowledgments

The author wishes to express his sincere appreciation to the following N.Y.C. projects and the supervisory personnel who provided the time and facilities necessary for collection of the data utilized in this study.

Mr. Theodore Brown
Monmouth Community Action
Program (N.Y.C.)
Longbranch, New Jersey

Mrs. Harriet Truman
Mercy Douglass Hospital
Philadelphia, Pennsylvania

Mr. William Morgan
United Progress, Inc. (N.Y.C.)
Trenton, New Jersey

Mr. John Rosario
Jersey City CAN DO (N.Y.C.)
Jersey City, New Jersey

Mr. John Bell
City of Paterson, New Jersey
Action Program (N.Y.C.)
Paterson, New Jersey.

Mr. James Heggins
Youth Opportunities Unlimited
Salisbury, Maryland

Mr. Norman Watts
Greater Chester Movement
Chester, Pennsylvania

Mr. Lester Wooten
Human Resources Agency
Manpower Operations
New York, New York

Mr. Fred Cahill
Nassau County Youth Board (N.Y.C.)
Minneola, New York

Mr. Thomas Brown
Atlantic City Human Resources,
Inc. (N.Y.C.)
Atlantic City, New Jersey

Mr. Edwin Hamilton
Macon County Board of Education
Tuskegee, Alabama

Thanks are also due the counselors associated with the above projects who were instrumental in the scheduling of testing sessions and in providing evaluation data for enrollees.

Assistance by Mr. Franklin Evans of ETS in administration of the measures and in analysis of the data is especially appreciated.

Table of Contents

Page

I. Introduction

A. Purpose of the Study 1
B. The Neighborhood Youth Corps Program 3
C. Program Goals and Objectives 8
D. Measurement Problems with Culturally Deprived Groups 12
E. Available Formal Measures 19
F. Evaluation vs. Prediction 22

II. Development of the Preliminary Measures

A. General Characteristics 25
B. Description of the Measures 29

III. Method

A. Sample 42
B. Administration 43
C. Data Analyses 46

IV. Results and Discussion

A. Item Analyses 48
B. Urban-Rural Comparison 50
C. Reliability and Validity of the Enrollee Measures 53
D. Intercorrelations and the Factor Pattern 58
E. The Criterion Scales 71
F. Discussion of Specific Enrollee Measures 75

V. Conclusions and Recommendations 91

Development of Evaluation Measures for Use with Neighborhood Youth Corps Enrollees

Abstract

Available formal measures of the paper-and-pencil variety have generally been criticized as inappropriate for appraisal of individuals from so-called culturally deprived or disadvantaged backgrounds. This is especially so when a range of intellectual and attitudinal variables are to be considered.

The present study represents an initial phase in the development and application of a battery of measures that would be suitable for research and evaluation with disadvantaged adolescents in the setting of a youth-work training program. The paper-and-pencil instruments devised attempt to compensate for deficiencies in currently available tests by shaping content, format and administration to the academic level and cultural background of the intended respondents.

A battery of 13 test booklets, incorporating such characteristics as relatively low verbal levels, oral presentation and in many cases pictorial formats, dealt with areas of vocational orientation (plans, aspirations, knowledge and interests); attitudes (toward self, authority and social values) and job-related reasoning skills. Based upon analyses of item characteristics, reliability estimates, validity (against proficiency ratings), factor analysis of the battery and research-related findings, a number of the measures were shown to possess a pattern of psychometric characteristics that warrants further research on their refinement and application.

Development of Evaluation Measures for Use with Neighborhood Youth Corps Enrollees

I. INTRODUCTION

A. Purpose of the Study

The importance of systematic evaluation in gauging the effectiveness of any broad program of social change seems axiomatic. However, for such programs, usable quantitative tools to assess and describe those changes that reflect desired goals have often been found lacking. Or, more likely, available measures are considered inadequate as a basis for judging many areas of program performance. This appears to be equally so for the class of social programs that fall under the rubric of "youth-work training" and especially for those categories of goals concerned with changes in the abilities, attitudes, or adjustment of the individuals served.

Sweeping criticism of youth-work training programs has been leveled by Herman and Sadofsky (1966) for a general failure to provide the scrutiny and potential modification of program components that derive from continual efforts in evaluation and research. Dubner (1965) is more specific in outlining the need for tests and measures of individuals in the many phases of a training program for adolescents from a poverty level culture. Only a superficial reading of the sections on evaluation, reported for hundreds of employment-training and work-study programs in the congressional report on "Profile of Youth - 1963" (U. S. Senate Committee on Labor and Public Welfare, 1964), provides dramatic evidence of a reliance on "testimonial" evaluation. There is a distinct avoidance of quantitative or systematic assessment of either the total program outcomes or of the effects of these training programs on the youth involved in them.

But, there may be some question of whether such criticism is fair and whether the expectations for definitive research and evaluation with youth-work programs are entirely realistic. Can good intentions be fulfilled if there are formidable limitations in the availability of measures which would be suitable for conducting the desired assessments? In the area of individual performance measurement, the question is highly relevant--particularly when one considers the population of adolescents for whom youth-work programs are intended and the characteristics of those formal assessment ("psychological") instruments that are currently available.

Problems arise, largely because of the specific content, format and standardization of these measures when applied to individuals whose somewhat unique cultural experiences, motivational patterns and skills are shaped by a poverty-level background and who, in many cases, are also minority group members. Various inadequacies in available tests, as utilized with members of minority or disadvantaged groups, have generated considerable professional concern (Campbell, 1964; New York City Commission on Human Rights, 1968; SPSSI, 1964). A strong implication of this concern is not only skepticism about existing test materials, but a possible need for different measures appropriately tailored in content, administration and interpretation for use with such groups. In considering the evaluation needs of a youth-work training program, with its many specific goals that involve changes in individual performance, any implication of a need for new approaches to measurement takes on added significance.

As the largest of the governmentally supported youth-work training programs, the Neighborhood Youth Corps (N.Y.C.) must obviously be concerned with the availability and application of individual assessment techniques in the conduct of research and evaluation efforts. The techniques must be relevant and reasonably accurate since the means by which program accomplishments can be demonstrated are certainly no better than the quality of the tools used to assess those accomplishments. The present study, in general, focuses upon problems of measurement and evaluation associated with adolescent school dropouts whose status is customarily defined as "disadvantaged" or "culturally deprived" and for whom some form of work experience and vocational counseling are being provided in the setting of a youth-work training program. More specifically, this report is intended as an initial or preliminary phase in the development of a battery of measures which would be particularly suitable for evaluation of male and female Neighborhood Youth Corps enrollees. These measures are to deal with characteristics of enrollee performance that reflect a variety of youth-work training program goals and might, therefore, be expected to change as a result of the individual's participation in an N.Y.C. project. Results of initial field-testing of the battery and examination of the appropriate psychometric characteristics of each measure are to be considered in recommending the ones most suitable for further development and use with the N.Y.C. program and, hopefully, for use with disadvantaged adolescent groups in general. For this study phase, it should be understood that only readily available and limited, concurrent criteria will be utilized as well as consistency estimates based upon single test administration.

B. The Neighborhood Youth Corps Program

(1) General Description: In order to understand the choice of variables and the design of the measures used in the present study, it is

necessary to clarify the overall scope and purposes of the Neighborhood Youth Corps program in carrying out its youth-work training function.

Very briefly, N.Y.C. operates a nationwide federal program administered by the U. S. Department of Labor and the Office of Economic Opportunity for the purpose of providing job experience along with vocational guidance or counseling for male and female adolescents, 14 to 21 years of age*, who come from families with poverty-level incomes. These adolescents may either be full-time students in an "In-School" program or high school dropouts in an "Out-of-School" program. Sponsors for N.Y.C. projects are primarily local community agencies or other nonprofit organizations which provide guidance services, develop jobs in local community organizations as well as their own organization--and assume overall responsibility for administrative services in the conduct of their project. (For a detailed summary of the program and its operation see U. S. Department of Labor, 1967.)

Of major relevance for any evaluation effort are a number of characteristics that would be expected in a program which is national in its scope and geographically widespread. These include not only substantial variability in the types of agencies that serve as sponsors, but also in the training and background of professional personnel who work with the enrollees; in the type and level of jobs available for the enrollee and in the effectiveness of on-the-job experience provided. For example, enrollee jobs may range quite widely from tasks that involve conservation work in a local park to those of an assistant in the electronics repair shop of a military base. Although each enrollee is paid at least the federal minimum wage by the N.Y.C. project, there may be variation in the number of weekly hours of work permitted and thus, in the weekly wage paid at any given project.

---

*The age group which the program served at the initiation of this study was the 16-21 age group. This was changed to include the younger age levels in July of 1967 and although the measures used in this study were designed with the original group in mind, the wider age range is not considered as critical to the potential application of these measures.

Variability in the composition of the population served also contributes significantly to any evaluation considerations. N.Y.C. enrollees include a variety of ethnic and racial groups from rural and urban areas who, for example, may be Caucasian, Negro, Puerto Rican, Mexican or American Indian. The differences, stemming from varied language and cultural backgrounds, are apparent in their effect upon an evaluation effort. More subtle variations in the enrollee population from project to project may be introduced over time by social changes that are not the direct result of project administration or overall operation of the N.Y.C. program. One such change that can have a direct effect upon the characteristics of a large proportion of males available for the Out-of-School program is that of military requirements and criteria for induction as defined by the Selective Service System. Other social changes involve the initiation and expansion of various governmental or private youth-training programs which can operate to draw off selective segments of the disadvantaged adolescent group and change the composition of the population available to the N.Y.C. projects. Changes in the local job market can have an obvious effect upon N.Y.C. enrollment and there are even variations from city to city in the acceptance of government anti-poverty efforts among minority group members. Such community attitudes can and probably do affect the recruitment and retention of enrollees. It is of importance that such program variations at least be understood and that any evaluation variables chosen cut across as many subclassifications, or program differences, as feasible in order to design a workable evaluation system--as well as to understand its limitations.

(2) Pertinent Enrollee Characteristics: Since the present study is concerned with the evaluation of adolescents enrolled in N.Y.C., some additional aspects of the relation of the enrollee to the program and his individual characteristics are important in shaping what is to be measured for evaluative purposes, and how this is to be accomplished.

An initial consideration is that the adolescents, with whom the results of this study deal, are school dropouts. Although the N.Y.C. In-School program is intended to serve those who are full-time students and to assist them in remaining in school, the disadvantaged adolescent group for which available tests and measures are least likely to be applicable is the one with the greatest "handicaps"--i.e., the unemployed dropout. Thus, the problems of evaluation, as will be shown in subsequent discussion, are most pressing for this group. A characteristic of the dropouts that is important in shaping evaluation instruments is their relatively low verbal level--particularly for males. Average reading level is found to be slightly above fifth grade for the male enrollees with the low point of the range at little more than second grade. The mean value for females is approximately a grade level higher.* The largest proportion of the enrollees are young adolescents, primarily in the 16-18 year age range, with males having left school at about the 10th grade--often for academic and/or disciplinary reasons and an expressed "lack of interest" in school. Most have never been employed in full-time positions before. Females have generally left school at about one grade level higher than males and less often for academic and disciplinary reasons than for financial reasons and requirements for child care. Their lack of job experience is found to be even greater than for the males (Tucker, 1966; Walther & Magnusson, 1967).

---

*These conclusions should be considered with caution since they are based upon limited evidence consisting of: (1) a sample of 100 male and 100 female scores on the Iowa Achievement Test administered at an Urban Youth Opportunity Center and (2) conversations with professional personnel at N.Y.C. projects and at a Youth Opportunity Center serving a major city. No adequate national sample of reading scores is known to be available for this particular adolescent population.

Enrollment in N.Y.C. is entirely voluntary on the part of the individual, as is the decision to remain. Enrollees may leave at anytime they choose, whether because of dissatisfaction with the program or simply because suitable full-time employment is found. No specific period for remaining in the program is mandated so that the time spent by an enrollee in an N.Y.C. project is highly variable. However, about six months is considered (very loosely) to be a "typical" maximum, although this limit may not be enforced by any given project. Thus, there are no firm criteria as a basis for defining completion--whether in terms of specific skills developed, employment obtained, or time spent in the program.

There are typically two adults with whom enrollees have their primary day-to-day contact and who assume major responsibility for the enrollee's progress during his stay in the program. One is considered a guidance counselor at the N.Y.C. project facility and the other is a work supervisor at a job site. The former individual is an employee of the N.Y.C. project and the latter is usually an employee of the agency which has agreed to accept the enrollee on a job. Counselors often maintain close liaison with work supervisors in order to deal with any common areas of concern regarding the enrollee and for purposes of evaluating his progress.

C. Program Goals and Objectives

Clearly, an evaluation system must reflect in some meaningful way the overall or long-range goals of the program under evaluation and, perhaps, a number of more immediate objectives that logically lead to those desired goals. In the case of N.Y.C., as for other youth-work training programs, the goals have tended to cluster under an "umbrella" criterion referred to as "employability" (U. S. Office of Economic Opportunity, 1967). Any outcomes of youth-work training efforts defined in terms that deal with obtaining and holding employment are certainly defensible. But if such outcomes are too narrowly defined and linked only to specific measures of enrollee employment following participation in N.Y.C. (e.g., time required to obtain employment, job tenure, salary, etc.), then the evaluation system risks placing too many eggs in one "criterion basket." Such overall performance criteria can fail to account for important changes in the individual enrollee which are brought about by N.Y.C. participation and linked to ultimate employability, and thus can also fail to identify those specific components of the program which need strengthening or modification. More explicitly, an N.Y.C. project may be carrying out its intended functions quite adequately. However, in its particular locale, the job market may be chronically poor, biases more evident against employment of minority group members, or the level of available jobs exceptionally low. If post-N.Y.C. employment criteria are the sole measures upon which evaluation is based, then obviously benefits reflecting individual change, as derived by the enrollee from the N.Y.C. project, are almost certain to be overlooked. In essence, there are advantages in assessing a variety of changes in individuals, as well as those broader program goals that are customarily used to define a criterion of employability.

Specific short-term program objectives that are reflected in changes in enrollee behavior constitute the basis for selection of the measurement variables considered in this study. Conversations with N.Y.C. program personnel (e.g., counselors, project directors), N.Y.C. headquarters personnel at U. S. Department of Labor and examination of the limited research on training programs for disadvantaged adolescents lead to reasonably good agreement in defining many of the objectives that are seen as important. Five such areas (which are recognized as having considerable overlap) can be defined grossly in order to provide a convenient framework for the content and development of specific measures.

These are summarized as follows:

(1) Vocational Orientation: This category of outcomes is concerned with the individual's general appreciation of the world of work. Job experience and vocational counseling are expected to provide him with a better awareness of what kinds of jobs might be available, the background, experience and training that these require and what a job involves in terms of its duties or demands on the worker. As part of this, there seems to be an attempt to develop realistic perceptions of jobs that would be reasonably attainable and to have the enrollee shape his future plans accordingly.

(2) Skills in Seeking and Maintaining Employment: A considerable amount of effort is apparently expended by counselors and work supervisors in instructing the enrollee on how to go about finding a job, and how to behave once he is employed. Such instruction may include where to look for jobs, how to make application, how to conduct oneself during an interview, proper dress or grooming on a job, and how to behave toward a supervisor and one's fellow employees--in essence, the rules of conduct generally accepted on a job; as amorphous as these may be in many instances.

(3) Attitudinal Factors: There is significant stress in youth-work programs placed upon developing "proper attitudes" on the part of the enrollee. These range from attitudes toward oneself to attitudes toward society in general, and especially toward authority figures encountered in the society. The rationale is based upon the obvious assumption that enhanced self-esteem is valuable for developing an individual's aspirations and for ultimate adjustment in the vocational as well as personal-social areas that, in addition, a certain degree of control of overt hostility toward authority is essential for maintaining employment, as well as for keeping out of trouble (primarily legal) in various other spheres of personal and social activity. A group of closely related attitudes that many counselors also seek to instill are a somewhat broad pattern of behaviors based upon a willingness to plan ahead. These are best typified by such activities as handling finances (e.g., saving money); making sacrifices for future education and recognizing the consequences of hasty aggression toward others and the legal difficulties that can follow. The pattern is one that sociologists have customarily defined as "deferred gratification" (Schneider & Lysgaard, 1953; Straus, 1962).

(4) Job Motivation and Interest Patterns: Another important but somewhat different facet of attitude development is geared toward shaping motivation of the enrollee in a way that leads him to see a possible "payoff" in working and in striving for job advancement. This may be reflected in a desire to go on for additional training that leads to better jobs, a willingness to put more effort into the job itself, in seeking employment or in conforming to rules on the job.

Since patterns of interest in job tasks are also believed related to job motivation, one of the intentions of a program that offers work-training experience and vocational guidance is to enhance motivation by helping the individual to define and clarify those interests. Thus, there is an assumption that, through experience and vocational information, the enrollee might crystallize or sharpen his vocational interests--perhaps even to "upgrade" the level of job tasks that form his interest pattern.

(5) Practical Reasoning in a Job Situation: This final category of objectives is probably stressed most frequently in a work-training situation by a job supervisor who expects the trainee to "use his head" or "stop and think" when carrying out some assigned task. At least, the intent is that he pay careful enough attention to directions and use reasonable judgment in carrying out work assignments. This form of practical judgment on the job undoubtedly encompasses many attitudinal and motivational components rather than being purely an aspect of what is usually considered "basic intelligence." Personnel in work-training programs certainly behave as if the characteristics that define these qualities can be instilled through training. Thus, it appears worth considering a set of job-related, reasoning skills which may possess unique dimensions that are subject to change as a result of job experience and proper guidance.

It should be apparent that the fairly cursory discussion of possible N.Y.C. program goals outlined here in no way exhausts a number of other possibilities, or denies their importance in youth-work training programs. Remedial training in arithmetic and verbal skills is undoubtedly an area of value as are any number of technical (vocational) skills that may be important to job performance. It is simply a fact that neither remedial skills instruction nor the formal development of specific technical abilities appear as uniform requirements for even a majority of the N.Y.C. projects.*

In any event, the sort of objectives outlined above are believed--by a number of professionals in the youth-work training field--to be at least as important to later job success, for economically and culturally deprived adolescents, as are specific remedial and technical skills. It is also an apparent assumption that these sorts of objectives are not only modifiable as a result of participation in N.Y.C., but important aspects of enrollee performance in that they can influence (i.e., are correlated with) post-program job adjustment or success. The remaining sections of this report are intended to translate these broad objectives into measurable terms and to demonstrate the utility of measurement devices that reflect such objectives.

D. Measurement Problems with Culturally Deprived Groups

Certain problem areas associated with measurement as it relates to disadvantaged adolescents (often dealt with in the context of minority group membership) must be taken into account, before suitable evaluation measures for this population can be chosen, designed, or utilized. Failure to do so is, in fact, the basis of much of the criticism aimed at currently available formal tests when applied to culturally deprived groups of any age level. The most recent interest in these problems of testing has derived from the industrial setting and involves questions regarding the applicability of available tests for employment selection purposes when applied to minority groups and/or culturally deprived applicants. Discussions of possible discriminatory effects of testing in personnel selection have been presented by Guion (1966, 1967) and by Kirkpatrick, Ewen, Barnett and Katzell (1967).

---

*This was unquestionably the case at the time this study was initiated and still remains largely true--although there has been an increase in the number of remedial skills programs undertaken by N.Y.C. projects.

The problems as viewed by industrial psychologists center around discriminatory bias of presently used measures, on the basis of their differential test validity (i.e., some tests may be valid for white individuals, but have little or no validity for Negroes). Some indication that this is a legitimate concern has been demonstrated by Kirkpatrick et al. (1967) and Lopez (1966). The types of solutions proposed for dealing with this form of test bias--given the existing measures--have involved separate analyses and interpretation of test results for different ethnic or cultural groups (e.g., use of different regression lines).

Much earlier recognition of unique evaluation and testing problems for individuals from economically and culturally deprived backgrounds has come from the field of educational evaluation. Formal techniques in educational appraisal of the abilities of disadvantaged children have focused almost exclusively upon intellectual or cognitive functioning, with continuing attempts made to overcome deficiencies in existing measures by resort to "culture fair" materials (Anastasi, 1965). These attempts to achieve culturally "neutral" content in measures of intellectual skill have been severely criticized and considerable skepticism remains regarding their utility in an educational setting. Wesman's (1968) discussion of some "follies committed" in attempts to develop culture-free instruments is especially relevant.

The point to be stressed, however, is not only that the pattern of educational testing for culturally deprived youngsters has been concerned almost exclusively with a search for intellectual indices, but as is the case for the industrial setting much of the research in this problem area has been dependent upon scores on existing intellectual or achievement measures to demonstrate differences between ethnic groups or to correlate test performance with degree of cultural deprivation (Coleman, 1966; Deutsch & Brown, 1964; Dreger & Miller, 1960). Paradoxically, the tests used are often questionable for application to disadvantaged groups in the first place. The finding of such differences is not particularly surprising (although to some extent worth demonstrating in order to confirm their existence). A more logical approach would appear to reside in recent attempts to determine why the differences exist in order to facilitate the education of culturally deprived youngsters, and to understand the test problems. There are findings, for example, of differences in background characteristics of the disadvantaged that result in different learning "styles," which in turn can affect test performance on a variety of different measures. Karp and Sigel (1965) in their review of research on testing disadvantaged children highlight what they term such "new approaches in appraisal" and cite suggestions by others (Gordon, 1963; Lennon, 1964) which carry strong implications that some new directions in test development are needed--not merely variations in interpretation of currently available measures. This implication is certainly apparent in the conclusion reached by Karp and Sigel (1965) that: "Psycho-educational appraisal of the disadvantaged pupil confronts us anew with the need to develop assessment procedures that both clarify the mechanisms by which learning occurs and guide the teaching-learning process."

Without unnecessarily belaboring an extensive literature, that has only tangential value for the immediate purposes of this study, the argument to be pursued here does not question the fact that the differences between culturally deprived and nondeprived groups do exist for a wide variety of formal psychological instruments and especially in certain cognitive-intellectual areas. Rather, it is that these measures are generally developed for middle-class individuals with their content, format, techniques of administration and normative data intended almost exclusively for such a group. Thus, results of testing with such measures merely confirm, in a broad sense, the very pattern of those differences that are a logical outcome of the cultural deprivation undergone by the disadvantaged. Recognition (or acceptance) of this assertion does not in itself provide a solution to the problem of usable measures for appraisal purposes--whether in meeting the evaluation needs of a youth-work training program or in an employment selection situation. Nor does reinterpretation of the results of existing measures for different groups necessarily circumvent or solve the problem. Any application of commonly used measures over the wide cultural and verbal range of the poverty level subpopulation, that is customarily defined as "disadvantaged," seems open to question. One wonders, for example, whether any customary interpretation of test results can be sensible when a formal paper and pencil verbal measure is applied in standardized fashion to an adolescent or adult whose reading level is near functional illiteracy and who completely misinterprets the content, or the specific intent, of a large number of the items.

The essential criticism of available measures and the rationale for considering new approaches might be stated as follows: any test obviously represents a sample of individual behavior and this behavior sample can be considered adequate only if the stimulus materials (i.e., test items) which elicit that response serve to communicate their meaning in reasonably clear fashion. Where such communication is essentially ineffective, the value of the response elicited is likely to be of limited applicability for any assessment purposes (e.g., no one argues against dropping "ambiguous" test items). In this light, a major conclusion that one can reach with regard to testing the culturally deprived is that there has been a failure to take account of their verbal skills or unique background and cultural characteristics in shaping test format, content and administration--so as to devise measures which, at the very least, communicate in the manner intended by those who utilize and interpret such measures (i.e., a 100 item attitude measure, entirely verbal in format, administered to an individual who cannot read a significant proportion of words and who lacks familiarity with much of the descriptive content may serve to verify the limitations of his particular cultural background, or verbal inadequacies--it is unlikely to be a measure of the attitude dimension(s) intended).

True, this is not the only problem found in testing minority and culturally deprived groups. The contention here is that it is a major one. Other specific aspects of background must also be considered in relation to tests and test-taking behavior before a suitable set of ground rules can be formulated for choosing or designing measures that would be applicable to a disadvantaged adolescent group. Continued trial and error with available measures, to see which of a vast number may be applicable for a wide variety of assessment needs, is not only laborious but may be largely fruitless. In light of even the limited knowledge of background characteristics, prior test performance and special problem areas in testing culturally deprived adolescents, one should be able to narrow the field and define relevant measurement techniques that have a reasonable likelihood of proving successful.

Those considerations, and their relation to tests and the measurement process, which should be taken into account include:

(a) A relatively low verbal reading level (as has been pointed out previously). This constitutes a major characteristic which limits the communication process and choice of test materials. It does not mean that the group is "non-verbal" in any sense. Even a third or fourth grade reading level provides a range of usable verbal skills--reading and oral--and does not necessarily imply that ideas must be communicated in an exclusively pictorial mode. However, existing measures designed for use with middle-class adolescents, that might also deal with suitable evaluation areas for a youth-work training program, are likely to possess too high a verbal level, whereas, measures having lower verbal level material are likely to be too childish in content for an adolescent.

(b) Performance of disadvantaged groups on intellectual ability measures is found to be superior in certain respects when oral presentation is used (Graham & Orr, 1966). Thus, comprehension for groups with low reading levels may be enhanced by utilizing oral communication along with any verbal printed materials.

(c) Differences in language meaning, or connotation, may exist for common words and phrases among minority group adolescents. Therefore, there is potential value in being aware of unique slang expressions and idiomatic modes of speech that can be highly meaningful to an adolescent or minority group culture--but not generally found in widespread use among middle-class adults. Similarly, in the area of communication and test content, there is the consideration of specific cultural experiences that are commonly found among middle-class youngsters but are relatively unfamiliar to a lower-class, minority group child. An obvious example involves choices of hobbies or avocational activities used as items in interest measures. These may be clearly understood by middle-class adolescents but because of economic consideration are often out of the realm of prior experience for the disadvantaged. Even with use of pictorial material in tests intended to minimize such criticism the content is often found to represent middle-class activities and to show individuals with strictly Caucasian racial characteristics (Geist, 1964; Weingarten, 1958).

(d) For the school dropout from a poverty background there is a complex pattern of negative attitudes toward academic trappings, of which formal testing is one. The general history of failure--and the individual's continuing awareness of it--is coupled with inadequacies in prior test performance on measures which have repeatedly exposed his poor verbal skills. There is little relation that such an individual can see between testing and what it means to him. It has, for the most part, served merely as an annoyance with no discernible "payoff." In fact, it may even have served as one of a number of barriers encountered in seeking employment.

(e) Since test-taking does not tend to be viewed favorably by disadvantaged adolescents, competition in such an area of performance is minimized--so that the individual does not always expend the effort to "do his best." Furthermore, much of the development of test-taking skills ("test-wiseness") stimulated in schools with predominantly middle-class white pupils (and by parental pressures) is minimal for culturally deprived and minority group pupils. The logical result of such experience is a response to testing that is often subsumed under "poor test-taking motivation." Lengthy tests with numerous verbal items that are not only poorly understood, but uninteresting in content, provide little to command attention and only contribute further to low motivation levels.

More subtle and complex problems of motivation can arise, which may influence test response by minority group members. Feelings about racial and/or class distinctions between examiner and examinee and particularly in the realm of conformity and response to an authority figure have been suggested as influences on test scores (Dreger & Miller, 1960).

(f) An effect upon test results, when low SES or disadvantaged groups are compared with middle-class youth, is the relatively lesser degree of consistency of response among the former group (SPSSI, 1964). Although this low reliability has been attributed to daily environmental factors that stem from economic deprivation, and lead to response instability, it is suggested that some of the variance may be accounted for by characteristics of the tests utilized and by content that stresses middle-class cultural experiences. The problem for consideration is one of whether items chosen as culturally relevant would reflect a greater range of meaningful variability than available formal measures have tended to uncover.

E. Available Formal Measures

Based upon an extensive number of measures that have been published, and others devised by researchers for utilization in psychological and sociological studies, it is legitimate to question whether any of these may be directly applicable to evaluation of disadvantaged adolescents in a youth-work training program. The number of available tests is immediately reduced to more manageable proportions if two critical requirements are imposed: first, that they be applicable specifically to the N.Y.C. program objectives and second that they be suitable for a population of culturally-deprived school dropouts. If in addition a third necessary requirement is imposed--that the instruments possess at least a reasonable number of the qualities customarily attributed to formal (usually published) measures--then the field narrows to a point of uncomfortably limited choice.

By formal measures (as contrasted with "informal" measures or questionnaires) it is generally meant that certain qualities exist in standardization of directions, administration and scoring; that there is information on statistical properties of items and total scores; that these psychometric characteristics show a pattern which meets acceptable "professional" criteria and, finally, that the measure is applied widely enough to provide normative data for the population of interest and, where necessary, for appropriate subgroups within that population. (For a general summary of formal and informal measurement see Greene, 1952, Ch. 2.)

Those available formal measures that might be considered to have possible applications in a group setting, for N.Y.C. evaluation purposes, often fail to be sufficiently compatible with the characteristics of the culturally deprived as discussed above. Among the reasons are, first, because they are found to be entirely verbal in format and at relatively high reading levels (in terms of items and directions). Those few that are pictorial measures are exclusively so and fail (for reasons unexplained) to enhance their information by combining verbal and pictorial material. In addition, the content (pictorial or verbal) is often culturally biased, dealing primarily with middle-class concepts and experience. Where low verbal-level material is used, the content reflects the experiences or interests of young children. Recognition of this limitation in purely intellectual measures has led one publisher to deal with the content of arithmetic and verbal skills items in a context that depends upon more common adult activities or experience--even though the material is at a low verbal reading level (Karlsen, Madden & Gardner, 1967).

Another major shortcoming is that the formal measures are often relatively lengthy (50 to 100 items or more for a measure). When combined with an entirely verbal format and poorly comprehended, or experientially irrelevant, content large numbers of items covering the same content area only add to the easy distractability and poor test-taking motivation attributed to culturally deprived school dropouts.

Additional contributions to the limitations of many formal measures are found in features that include tight time restrictions (especially when applied with noncompetitive individuals who are not "test-wise"), numerous items arranged on a sheet where sequence may be confused or where items are easily skipped over and, finally, errors engendered by even the customary process of transposing a response from the item choices on a page to a separate answer sheet (Clark, 1968). These and similar requirements, normally desirable for ease in the administration and scoring of formal group measures, lead to a remoteness of the examiner from the individual respondent in the test-taking process. For disadvantaged adolescents who will often fail to ask for clarification or assistance where appropriate, such highly structured administration may add unnecessarily to errors of measurement.

In the area of informal measurement, which derives largely from research with disadvantaged groups (more often in terms of a low SES category), the instruments used are usually drawn up by a researcher for his particular study. A considerable number of such informal measures (especially noncognitive ones) that have been used with disadvantaged groups are readily found in the literature. Although on the surface many such paper and pencil measures used for research studies appear applicable, their major drawback is, invariably, that by no stretch of the concept can they be considered "measures." They are more accurately termed "questionnaires," or a compilation of questions which are assigned scale values, hopefully with some ordinal properties, but with little if any analysis performed to indicate their measurement characteristics. Study results and conclusions are then reported as if measures of known characteristics and proven capability have been utilized. Any of the inconsistency that already exists in research results concerning disadvantaged adolescents can only be magnified by a failure to develop measures with known properties that are consistently measuring reasonably understood dimensions of behavior.

Although there were shortcomings found in both available formal and informal measures which would preclude their direct application as evaluation indices for disadvantaged youth, a number of these nevertheless provided important clues to content and item format that were incorporated in a number of the measures developed for this study. A survey of many of these applicable test materials is reported in the draft of a specifications document prepared in the course of developing the measures (Educational Testing Service, 1967). Some portion of this background material is summarized in the description of the measures that appears in Section II of this report.

F. Evaluation vs. Prediction

In developing any form of behavioral measure, an appreciation of its intended use(s) bears emphasizing. There can be important implications for data analyses and the decisions reached in test construction, when tests are considered for single or multipurpose application. The issues involved in test utilization as a function of measurement properties are fairly complex and arguable. In order to confine this brief discussion to some of the more pertinent issues for the present study, the assumption made is that the primary purpose of the measures to be developed here is one of evaluating certain N.Y.C. enrollee characteristics over the period of months that the individual remains in the program. Given this assumption, it follows that if the numerous program objectives for N.Y.C., as outlined above, could be translated into highly precise forms of enrollee behavior--that are readily observed or measured--then the items comprising a given test could be formulated as direct statements of such behavior or performance. Any measure containing items that define such specific aspects of performance would be considered valid on the face of its content ("face validity") and it might be argued that there should be little concern with item characteristics, or test reliability and predictive validity. Face validity invoked for evaluation purposes merely says that the enrollee's test reponses, after some defined degree of participation in N.Y.C., are expected to reach given levels and the closer he comes to reaching these defined standards the greater the success of the program. Thus, the evaluation of change represents, in effect, shifts toward or away from the desired behavior. This sort of argument for "representativeness of content" as the primary basis for design of an evaluation measure has been advanced by Horn (1968) and serves to minimize reliability (or predictive validity) as a consideration.

In practice, unfortunately, one is rarely lucky enough (or gifted enough) to be able to define test content and desired individual responses so clearly--least of all in a social program that has rather broad goals and subgoals involving individual improvement. There is simply not enough known about the specific individual behaviors to be changed (or the form of test items that would directly reflect such behaviors). And, more important, there may be nagging suspicions that what are thought to be desirable outcomes in enrollee behavior at the completion of the N.Y.C. program may be relatively unrelated to longer range (i.e., post program) goals. Where this possibility exists as it usually does--and the N.Y.C. program would appear to be no exception--there must be some attention paid to reliability of the measures (or its estimate based upon internal consistency) as well as some aspects of predictive validity. Also as Ebel (1968) has pointed out in defense of such a position, tests intended for purely evaluative purposes often end up being used for decisions that affect individuals, for example, by an N.Y.C. counselor who utilizes them for guidance purposes or for assignment of enrollees to job sites.

The point is that it can be reasonable to have a multipurpose test battery in which measures that serve an evaluative function may also serve, to some extent, a successful predictive or classification purpose. Where the tests possess at least acceptable levels of reliability, this is possible and it is, in fact, customary for educational tests to serve both purposes (Ebel, 1968).

However, although consideration will be given to estimates of reliability and predictive validity for the evaluation measures under consideration, it can be self-defeating to force such issues too far in expectations for relatively high reliability (or validity) coefficients at the stage of test development represented by the present study. One reason for caution in dealing with reliability is a practical paradox in requirements for reliability of evaluation instruments that are intended to be sensitive to individual change over time. If high reliability implies test-retest stability over any extended period, then a highly reliable measure of some particular characteristic might be useless for a program that intends to effect individual change over a period of months and to evaluate such change. Secondly, there is concern with the issue of reliability attenuation resulting from the previously cited group-homogeneity for low SES or culturally deprived. Finally, there is the possible effect on reliability that stems from constraints on test length. The desirability of shortened test length follows from the prior discussions of test-taking problems among disadvantaged adolescents. Therefore, unusual caution is needed before applying any stringent reliability demands in the choice of measures intended for the assessment purposes and the particular population to be evaluated here.

II. DEVELOPMENT OF THE PRELIMINARY MEASURES

A. General Characteristics

On the basis of those performance objectives that can be defined for the Neighborhood Youth Corps Program, the characteristics of the enrollees who are to be evaluated and the admittedly sketchy knowledge that exists on problems of assessing the capabilities of disadvantaged adolescents, it is possible to define some general ground rules that should be met by a battery of relevant evaluation measures; as well as their specific content. At the outset, it should be evident that any measures dealt with, at this phase of development, cannot be considered as more than preliminary in nature. They can certainly not be construed as meeting the requirements of formal instruments.

For the battery of 13 enrollee measures prepared for use in this study, the following guidelines were applied:

(1) Format: Test items are printed in small booklets (approximately 5" x 8") with the content for each item generally appearing on only a single page. (For about 25% of the test material there are as many as three items spread over two facing pages of the booklet.) Seven of the 13 test booklets contain pictorial information, accompanied by supplementary verbal material below, or within, the pictorial scene. The intent is to communicate as much information as possible by not relying exclusively upon either a pictorial or verbal mode. Boxes to be marked by the respondent are next to each item-response choice on the page. (A glance at the sample items in the pages of this section provide an overview of the format.) Verbal directions which serve as a capsule summary of the longer, orally presented directions are printed on the front cover of each test booklet, where space is also provided for the individual to write his name and N.Y.C. project. No test titles appear on the cover page, or anywhere else in the booklet. Separate male and female forms are required for the seven measures that utilize pictorial information, but the format always remains the same despite appropriate content changes.

This format requires a considerable amount of art work, more complex printing than is customary, and hand scoring of items. Thus, it is apparent that aspects of logistical convenience in test development and usage were intentionally ignored for this phase of study.

(2) Verbal Content: Printed material (which is also orally presented) was maintained at a relatively low verbal reading level. Approximately a fifth grade reading level was chosen and verified by a check of the test booklet having the most difficult reading level against the Thorndike-Lorge Word List (Thorndike & Lorge, 1944). (However, an estimate of comprehension level for orally presented material, that is also printed, cannot be made with any precision.) Directions, both oral and printed, are kept at a similarly low verbal level. Slang expressions, generally suitable for adolescents, are used where they seem to fit--but not stressed unnecessarily in the overall verbal style. Verbal content and concepts are, insofar as possible, relevant to and understood by adolescents from a disadvantaged or low SES background.

(3) Pictorial Content: All pictorial scenes depict a youthful (adolescent) figure in some situation that portrays the item intent for the given measure. Thus, clarity of pictorial content is critical and simple black-line drawings (i.e., "cartoon" style) were utilized. Since verbal material is coupled with the pictorial material, the information in either mode tends to be supplementary so that ambiguity can be minimized.

Any choice of pictorial material to accompany the verbal was always made on the basis of its providing some degree of supplemental information and not merely for uniformity or for its own sake. Pictorial information is also intended to provide some degree of added interest to the test content and, hopefully, to enhance motivation (Weingarten, 1958). If overdone, however, with little apparent relevance, its use can appear somewhat demeaning to an adolescent--despite his low verbal reading level. Where the measures are to be applicable to a number of racial or ethnic groups and the respondent is to identify with the youthful figure in a pictorial scene, it seems logical not to portray any particular racial characteristics. However, this is rather difficult for an artist to portray and the choice was made to "neutralize" (or at least minimize) any features that depict race by specifying simple line drawings of the type that are found in the Rosenzweig Picture Frustration Test (Rosenzweig, 1948). The scenes depicted are intended to possess some degree of situational familiarity to the individual. Although this is difficult to accomplish in picturing a number of job-tasks and job-related scenes, many of the pictorial scenes are probably part of a broad enough frame of cultural reference on the basis of their appearance in printed media, in movies, or on television.

(4) Administration: The measures are intended for administration to small groups (not exceeding about 12 individuals per group), with separate sessions for testing males and females. Directions as well as all item stems and response choices are orally presented by the examiner. Thus, there is a specific intention to minimize the influence of any verbal reading component on each measure. As will be apparent in the detailed description of the measures, a small porportion of the items do contain verbal material accompanying the item stem and response choices (e.g., a newspaper want ad) that could not logically be read by the examiner without immediately revealing the correct item responses.

(5) Test Length: Since the "test-taking-attention-span" of the intended respondents can be expected to be relatively short, it is prudent to maintain test length at the minimum thought necessary to cover a desired dimension, and to do so with a sufficient range of content to result in a somewhat reliable measure. In addition, at the present stage of development,

and for the relatively unique approach to test material and format used here, it would be wasteful to construct numerous items for any measure in which all the items of a given type might prove useless.

B. Description of the Measures

The 13 preliminary evaluation measures for use with male and female enrollees are to be described in this section, along with two criterion rating scales used by N.Y.C. counselors and work supervisors in evaluating the enrollee. Brief discussions of the content, format and scoring, as well as some aspects of the rationale, are presented along with samples of items from each measure. This material is summarized from more detailed discussions of these areas, available in the previously cited document on specifications for the measures (Educational Testing Service, 1967).

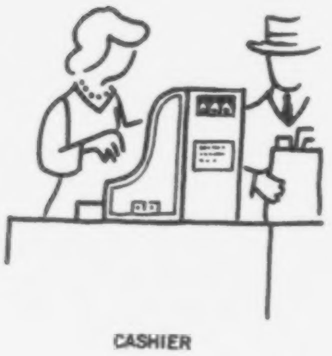
Job Knowledge (JK)

This measure, consisting of 27 items for the female form and 30 items in the male form, is intended to assess the degree to which the enrollee is aware of a number of categories of job requirements. For each form there are three multiple choice questions asked. These relate to an occupation pictured, along with its corresponding job title printed below. In the 30-item male form there are five items covering each of the six job knowledge categories--(1) education required, (2) where the work is performed, (3) full-time starting salary, (4) the specific major task performed, (5) the most likely working hours required and (6) tools required (with response choices consisting of pictured tools). The female form contains only the first five knowledge categories since female occupations for which tools can be identified are difficult to uncover.

No published, formal, measure is known to exist in the area of job knowledge that would be suitable for this battery. The "Vocational Interest and Sophistication Assessment" (Parnicky, Kahn, & Burdett, 1963) that has been under development for use with individuals having low verbal reading levels is used to obtain scores on both "knowledge of job conditions" and vocational interest." But it is exclusively pictorial and was designed for use specifically with educable mentally retarded. Thus, although its format and rationale for pictorial material served as a valuable aid, its direct application to a culturally-deprived adolescent group was considered questionable.

Scoring of the Job Knowledge measure is based upon total number of correct responses obtained from the four-choice, multiple-choice, items and can also be broken down into subscores based upon the number of correct responses for each of the Job Knowledge categories.

Figure 1A

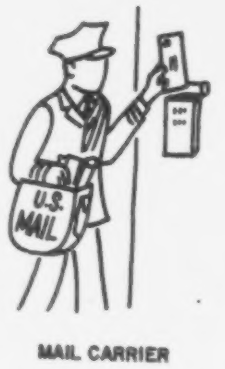


Where would you work on this job most of the time?

☐ In a factory on the assembly line
☐ In the kitchen of a restaurant
☐ In a store
☐ In your own home

Copyright © 1967 by Educational Testing Service

Figure 1B



About how much would you get paid when you start on this job full-time?

☐ $2.25 to $2.75 an hour
☐ 50¢ to 75¢ an hour
☐ $1.00 to $1.50 an hour
☐ $3.50 to $4.50 an hour

Copyright © 1967 by Educational Testing Service

Vocational Plans and Vocational Aspirations (VA, VP)

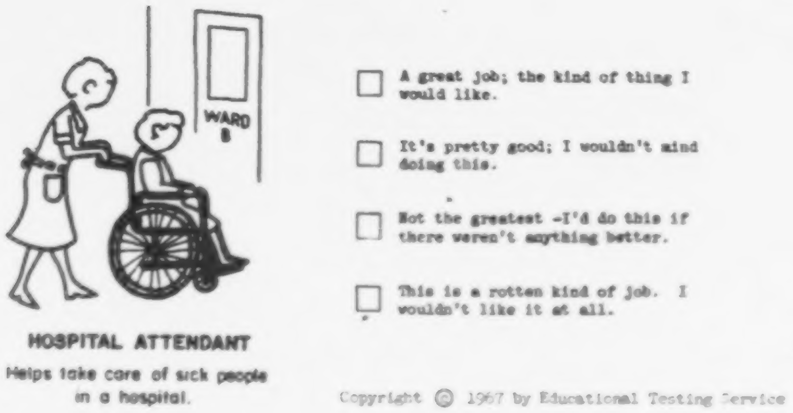
These two measures are appropriately considered together since they utilize the same stimulus materials (although different response choices) and are also scored in relation to one another. Each booklet contains, in both male and female forms, 16 items in the format shown below with pictorial and verbal material for each item on a single page. The items were chosen on the basis of their social status levels. Status values were derived from the NORC occupational prestige rating studies (Duncan, Hatt, & North, 1961) and the 16 items have been grouped into eight status levels based upon these values, with two items at each level.

The individual responds to each occupation shown on a four-point scale that indicates the degree to which he plans (or aspires) to that occupation. There are no formal published scales for measurement of vocational plans or aspirations, as intended for use in this study. Two published measures which deal with occupational stimulus materials are, in fact, used for personality assessment purposes (Ammons, Butler, & Herzig, 1949; Holland, 1959). Highly informal check lists and questionnaires, or simply open-ended questions, have been used in research with adolescents or young adults (Empey, 1956; Ginzberg, Ginsburg, Axelrad, & Herma, 1951; Slotkin & Forlano, 1962; Stephenson, 1955; Varo, 1964). But the nature of their content, format, scoring and lack of measurement properties would make them inappropriate as applicable formal instruments.

Scoring for the VA and VP measures consists of a "status discrepancy" score for each one, based upon the difference between the weights assigned by the respondent to each status pair. (For example, if the pair of items at status level eight (8)--"physician" and "scientist"--were assigned degrees of choice "4" and "2" respectively, the status discrepancy at that status level is a "2.") The total discrepancy score is thus the absolute sum of these individual values for all items in the measure. In effect, this is intended as a measure of the consistency with which the individual uses an underlying status continuum for his vocational choices. A second score is possible by utilizing the difference between the level (i.e., degree) of choice assigned an occupation in terms of plans and aspirations. For example, "Auto Mechanic" may be assigned a high degree of aspiration with a weight of "4" but only a moderate weight (e.g., a "2") for the plans measure. The difference score is thus the difference in these weights assigned by the respondent and summed over the corresponding occupational pairs in each of the two measures. This difference between plans and aspirations for the same occupations is felt to have possible uses as a measure of the degree of "reality" in vocational choice--based upon what the enrollee would like to do vocationally and what he expects to do.

Vocational Aspirations

Figure 2A



☐ A great job; the kind of thing I would like.

☐ It's pretty good; I wouldn't mind doing this.

☐ Not the greatest -I'd do this if there weren't anything better.

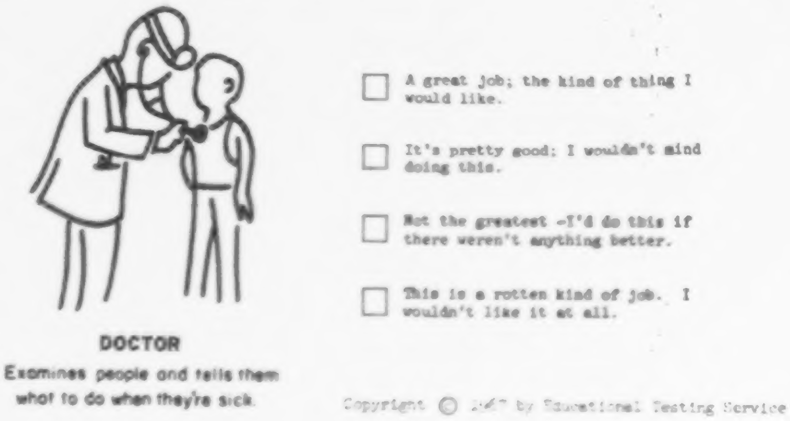
☐ This is a rotten kind of job. I wouldn't like it at all.

HOSPITAL ATTENDANT

Helps take care of sick people in a hospital.

Copyright © 1967 by Educational Testing Service

Figure 2B



☐ A great job; the kind of thing I would like.

☐ It's pretty good; I wouldn't mind doing this.

☐ Not the greatest -I'd do this if there weren't anything better.

☐ This is a rotten kind of job. I wouldn't like it at all.

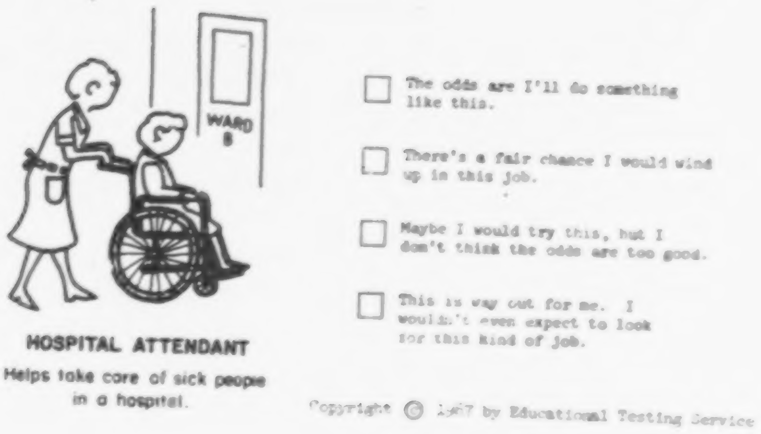
DOCTOR

Examines people and tells them what to do when they're sick.

Copyright © 1967 by Educational Testing Service

Vocational Plans

Figure 2C



☐ The odds are I'll do something like this.

☐ There's a fair chance I would wind up in this job.

☐ Maybe I would try this, but I don't think the odds are too good.

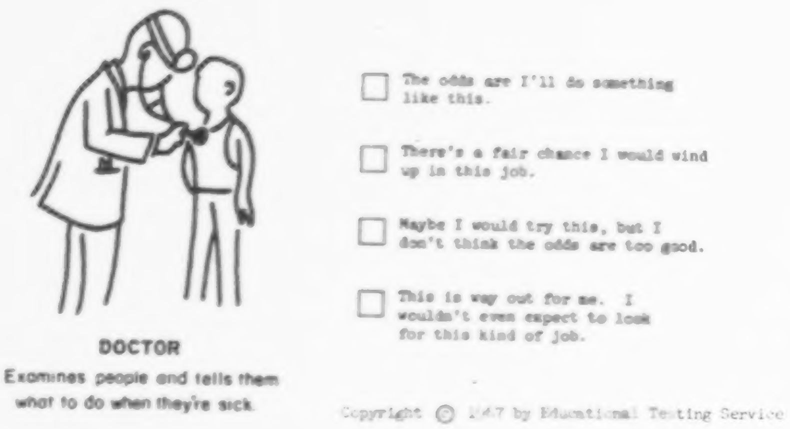
☐ This is way out for me. I wouldn't even expect to look for this kind of job.

HOSPITAL ATTENDANT

Helps take care of sick people in a hospital.

Copyright © 1967 by Educational Testing Service

Figure 2D



☐ The odds are I'll do something like this.

☐ There's a fair chance I would wind up in this job.

☐ Maybe I would try this, but I don't think the odds are too good.

☐ This is way out for me. I wouldn't even expect to look for this kind of job.

DOCTOR

Examines people and tells them what to do when they're sick.

Copyright © 1967 by Educational Testing Service

Interest in Vocational Tasks (VI)

As the title implies, the attempt here is to measure an interest pattern based exclusively upon specific occupational tasks, rather than interest patterns derived from choices based upon occupational titles or from the more customary combination found in formal interest measures of leisure time activities (i.e., hobbies) and vocational tasks.

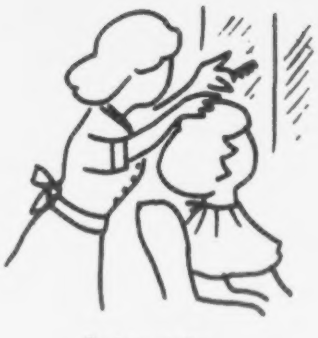
Criticisms involving verbal level, test length, techniques of administration and cultural relevance of content can be made of the widely used Strong Vocational Interest Blank (Strong, 1945) and the Kuder Preference Record (Kuder, 1956) when these measures are applied to a disadvantaged group. Although it was developed for men in nonprofessional occupations (i.e., most skilled and semiskilled), similar criticisms can be aimed at the Minnesota Vocational Interest Inventory (Clark & Campbell, 1965). Some of these criticisms have been touched upon by Clark (1961) and Barnette and McCall (1964).

In contrast to these verbal measures, there are two exclusively pictorial interest tests (Geist, 1964; Weingarten, 1958) which also possess shortcomings for use with disadvantaged or minority groups. The level of vocational tasks portrayed, the clarity of the pictorial material, the clearly Caucasian figures shown, and the standardization data can serve to limit their applicability. One of these measures, the Geist Picture Interest Inventory (Geist, 1964), has also been criticized on the basis of shortcomings in its development (Shimberg, 1965).

The approach to interest measurement that proved helpful in the design of the present measure is the 168-item, all verbal, California Occupational Preference Survey (Knapp, Grant, & Demos, 1966). Obviously, it is not a measure that meets the ground rules for direct application to the adolescent group of interest in this study. However, its technical approach and content serve as a valuable model, since it utilizes only specific occupational tasks extracted from the Dictionary of Occupational Titles (USES, 1965) and it scores the degree of interest in these tasks for occupational groupings or clusters that are derived from factor analyses of the items. By utilizing those items that could be pictorialized reasonably and those occupational groupings in which these items fell (along with some rewording of tasks where verbal level was considered too high), a 28-item measure was constructed with seven occupational clusters containing four items each. The seven clusters are termed: Clerical, Service, Technical, Business, Outdoor, Science, and Aesthetic.

Scoring consists of summing the weights assigned to the vocational tasks (1 to 4 along a "favorable-unfavorable" continuum) for the group of tasks in each of the seven occupational clusters.

Figure 3A



Do hair styling.

☐ Pretty bad - I couldn't take doing it.

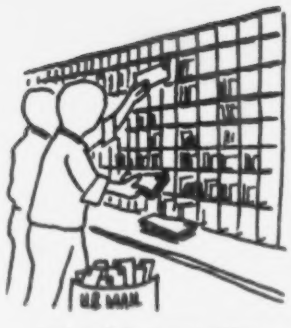
☐ Not too good - I might do this but I wouldn't like it.

☐ This is OK - I wouldn't mind this too much.

☐ This is great - just the kind of thing I would like to do.

Copyright © 1967 by Educational Testing Service

Figure 3B



Sort mail in a post office.

☐ Pretty bad - I couldn't take doing it.

☐ Not too good - I might do this but I wouldn't like it.

☐ This is OK - I wouldn't mind this too much.

☐ This is great - just the kind of thing I would like to do.

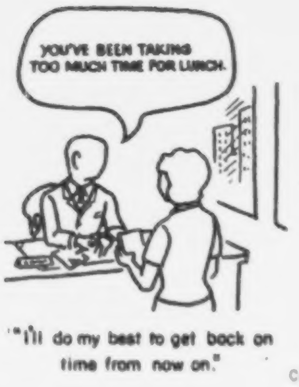
Copyright © 1967 by Educational Testing Service

Attitude Toward Authority (AA)

This attitudinal measure consists of 12 items in the female form and 13 in the male form. The single item on a page always depicts an adolescent figure being spoken to by an authority figure (i.e., judge, supervisor, teacher, policeman, father, mother, truant officer, job interviewer). The response given by the adolescent to the authority figure is either a clearly anti- or pro-authority response. Item stems are always in the same form as shown in the two sample items below--i.e., the number of people who you think would give that answer in this situation--with the five item choices covering some range between the odds of "1 out of 10" to "9 out of 10." The technique used here has come to be known generally as the "disguised factual" approach in testing and has been applied successfully to adults from a wide range of SES levels by Bernberg (1954). In his manual for the Human Relations Inventory (HRI), Bernberg refers to this item style as the "direction of perception" technique and he has used it for measuring social conformity with the 39-item all-verbal HRI. Two other formal measures related to the conformity area, and usable with young adults, are the Socialization Scale of the California Personality Inventory (Gough, 1957, 1960) and the Family Opinion Survey (Elias, 1952) that measures attitude toward authority relationships within the family. These three formal measures were considered too narrow in content (and validation) and too highly verbal to meet the intention of the present measure, or to be adopted directly.

Scoring of the AA items is based upon the assignment of weights 1 to 5 along a continuum of choices with a weight of 5 representing the high, or proauthority, end of the scale. Thus, for example, if the item depicts an antiauthority response made by the youth to an authority figure, a response choice of "1 out of 10 people" would receive a weight of "5" (the respondent is proauthority and thinks that few people would make such an antiauthority remark); while a response at the other end of the continuum ("5 out of 10 people") on that same item represents an antiauthority attitude and receives a weight of "1." Based upon Bernberg's findings, it makes no difference in the overall score whether the response-choice continuum on any item ranges in steps from, say, "1 out of 10 people" to "5 out of 10 people" or from "5 out of 10 people" to "9 out of 10 people." It is only the relative location on the continuum that determines the contribution of the response.

Figure 4A

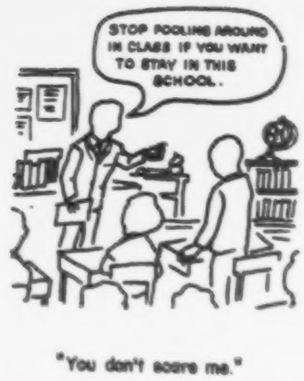

YOU'VE BEEN TAKING TOO MUCH TIME FOR LUNCH.
"I'll do my best to get back on time from now on."

Studies show that how many people give that answer to the boss?

☐ 5 out of 10 people
☐ 6 out of 10 people
☐ 7 out of 10 people
☐ 8 out of 10 people
☐ 9 out of 10 people

Copyright © 1967 by Educational Testing Service

Figure 4B


STOP FOOLING AROUND IN CLASS IF YOU WANT TO STAY IN THIS SCHOOL.
"You don't scare me."

How many students do you think would give this answer to a teacher?

☐ 2 out of 10 students
☐ 3 out of 10 students
☐ 4 out of 10 students
☐ 5 out of 10 students
☐ 6 out of 10 students

Copyright © 1967 by Educational Testing Service

Self-Esteem (SE)

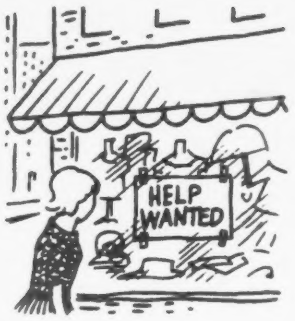
The 16 items that constitute this measure consist of pictorial scenes in which an adolescent is portrayed in various academic, social, and employment settings. The respondent is asked to imagine that the young person in the picture represents himself. The response required is one involving his estimate of what would happen if he were, in fact, the youthful individual in that scene and is intended to reflect his level of self-worth or self-esteem on a 3-point scale.

Each item (and the aggregate of items) are intended to deal with a dimension often conceived of as competence in interpersonal social action. Wylie (1961), who reviews much of the research literature in this area and the different approaches to quantifying self-esteem, notes the general lack of validity data to support the variety of definitions and measures used in dealing with the concept. As her review indicates, most self-esteem measures are informal, exclusively verbal and involve a fairly complex clinical rationale in their interpretation and scoring. One such informal scale that has been utilized on reasonable size samples of adolescents (but of unquestionably higher academic levels than an N.Y.C. dropout group) is a purely verbal 10-item measure that lacks adequate analysis of its psychometric characteristics; with item wording that is questionable for the N.Y.C. enrollee group (Rosenberg, 1965).

Although two published formal measures exist that provide some form of self-regard score (Gough, 1957; Rogers, 1961), there is no evidence of their suitability for a disadvantaged adolescent group. In addition, one of these--the Rogers "Personal Adjustment Inventory" (Rogers, 1961)--is intended for grade school children 9 to 13 years of age and its content is inappropriate for adaptation to older adolescents. The other measure is questionable for direct application in this study because of its verbal level and its general cultural content.

Scoring of the total SE measure consists of summing all item weights where the weights from 1 to 3 on each item represent a low-to-high continuum of self-esteem.

Figure 5A



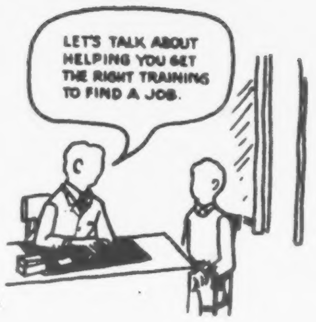
☐ I might try for a job in that store, but they probably wouldn't want me.

☐ They would turn me down cold and wouldn't think I'm worth giving a chance.

☐ They would think I'm worth hiring for a job in that store.

Copyright © 1967 by Educational Testing Service

Figure 5B


LET'S TALK ABOUT HELPING YOU GET THE RIGHT TRAINING TO FIND A JOB.

☐ There's probably nothing much he could do for me.

☐ I know I could make good on a job if he helped me get the right training.

☐ I might stand some chance on a job if I could get some kind of training.

Copyright © 1967 by Educational Testing Service

Deferred Gratification (DG)

The ability to delay present rewards for future gain (i.e., defer gratification) represents the intended attitudinal dimension for this 16-item scale. The booklet for this measure contains a total of 20 items, four of which are common slogans that serve as distractor items and are unscored.

The deferment of impulse gratification as a behavioral concept was introduced into the sociological literature by Davis and Dollard (1940) in an attempt to deal with patterns of behavior that differentiated the middle and lower social classes. More recent research dealing with measurement of this concept has been made by Schneider and Lysgaard (1953) and Straus (1962) who have attempted to quantify a "deferred gratification pattern" (DGP). Their results led to several scales that may have value for defining important predictive responses in adolescents. Straus (1962), for example, found modest correlations between the DGP and measures of academic achievement and occupational aspiration; even with intelligence and SES partialed out. Thus, this type of measure may be capable of meaningfully differentiating between individuals in the adolescent group that is utilized in the present study. The subscales with the greatest utility, based upon the prior research, appear to be those defined as: "Hasty Aggression" (e.g., "If a guy gets you sore, hit him--no sense in waiting until you calm down"); "Freedom of Spending" (e.g., "Money is to spend, not to save") and "Affiliation" (e.g., "You have to be careful about the kind of people you pick for friends"). A fourth subscale was derived from the more recent attempts to deal with gratification deferment in terms of specific delay of reinforcement as used in research with children (Mischel, 1961). This subscale is termed "Specific Delay of Reward" (e.g., "I would rather have a cheaper used car today, than save up until I could afford a new one") and is not one that has been used in the past as part of the DGP scales. Item-response format is similar to the disguised factual approach in that the individual responds on the basis of the degree to which he thinks other people might agree or disagree with the statement (5-point scale).

Scoring consists of the sum of the weighted scores over all items, with a high score assigned to a high level of deferred gratification.

Figure 6A

I WOULD TELL A COP OFF IF HE "BUGGED" ME - EVEN IF IT MEANT GETTING IN TROUBLE.

☐ Almost nobody would feel this way.
☐ Only a few people might agree with this.
☐ Some people would agree with this
☐ A lot of people would feel this way.
☐ Most everybody would agree with this.

Copyright © 1967 by Educational Testing Service

Figure 6B

BUY WHATEVER YOU WANT, THERE'S ALWAYS SOME WAY TO PAY IT OFF.

☐ Almost nobody would feel this way.
☐ Only a few people might agree with this.
☐ Some people would agree with this.
☐ A lot of people would feel this way.
☐ Most everybody would agree with this.

Copyright © 1967 by Educational Testing Service

Job Seeking and Job Holding Skills (JS;JH)

This measure is an attempt to deal with what are probably the most obvious areas of behavior and knowledge that it is felt an enrollee should develop in a youth-work training program--i.e., knowledge of what to do to find a job and, having done so, how to behave in order to retain it. The items used in the measure represent a direct translation of those areas of job-holding and job-seeking skills that can be defined reasonably, and are based upon the rather limited knowledge of just what these specific skills or behaviors might be. The limitations stem from a glaring lack of meaningful research concerning reasons for hiring and dismissal of employees and what an employer uses, as the most relevant considerations, to make such judgments. (Reasons that often range from outright bias to "pet" annoyances, or unwritten personnel policies, are not readily obtained from personnel files of a company.)

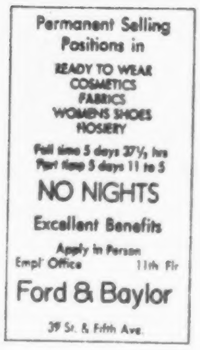
Nevertheless, some broad areas of agreement are possible on the most important areas of job-related behaviors and can be found in courses that have been drawn up for the purpose of teaching such skills (Educational Design Inc., 1966; U. S. Office of Economic Opportunity, 1966) or from the "how to" books or pamphlets that deal with job success (Feingold & List, 1958; Magoun, 1959; U. S. Dept. of Labor, 1964). No published formal measures of these job-behavior skills have been found in the test literature.

The first portion of the JS;JH measure consists of 17 items in a four-choice, multiple-choice format, that deal with job seeking skills (JS). Questions are concerned with where to look for a job, the contents of several want-ads, that are reproduced in the booklet, and questions about parts of application blanks. The respondent is required to choose the correct answer for each item. The second section contains 11 items dealing with on-the-job (job-holding) behavior (JH), covering the customary concerns with punctuality, appropriate dress, response to a supervisor's request and some items dealing with what are probably less-often revealed sources of employer discontent, such as placing bets and drinking on the job. The respondent chooses one of three response statements that describes how he feels--or what he would say, or do, in the particular situation shown. These responses are scaled "1" to "3" on a "low" to "high" (least correct to most correct) job-holding skills continuum.

Scoring for the 17 items of the job-seeking skills portion consists of the number of correct responses. For the 11 job-holding items, the response weights 1 to 3 on each item are summed over all 11 items. The two scores are always treated separately for purposes of analysis.

Job Seeking

Figure 7A



Permanent Selling
Positions in
READY TO WEAR
COSMETICS
FABRICS
WOMENS SHOES
HOSIERY
Full time 5 days 37½ hrs
Part time 5 days 11 to 5
NO NIGHTS
Excellent Benefits
Apply in Person
Empl' Office 11th Flr
Ford & Baylor
39 St. & Fifth Ave.

When would you work on this job?

☐ At nights only for five days a week.

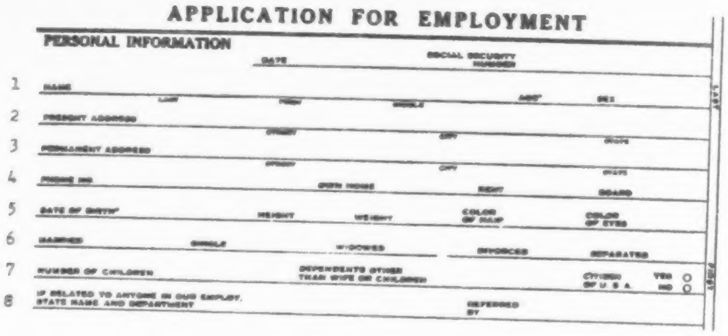
☐ Only on Thursday, Friday and Saturday.

☐ Only during the days.

☐ Any hours you want to pick.

Copyright © 1967 by Educational Testing Service

Figure 7B



APPLICATION FOR EMPLOYMENT

PERSONAL INFORMATION

DATE SOCIAL SECURITY NUMBER
1 NAME LAST FIRST MIDDLE AGE SEX
2 PRESENT ADDRESS STREET CITY STATE
3 PERMANENT ADDRESS STREET CITY STATE
4 PHONE NO OWN HOME RENT BOARD
5 DATE OF BIRTH HEIGHT WEIGHT COLOR OF HAIR COLOR OF EYES
6 MARRIED SINGLE WIDOWED DIVORCED SEPARATED
7 NUMBER OF CHILDREN DEPENDENTS OTHER THAN WIFE OR CHILDREN CITIZEN OF U.S.A. YES ○ NO ○
8 IF RELATED TO ANYONE IN OUR EMPLOY, STATE NAME AND DEPARTMENT REFERRED BY

Copyright © 1967 by Educational Testing Service

If you've never been married before which one would you mark on line number 6?

☐ Single

☐ Divorced

☐ Married

☐ Widowed

What would you put on line number 2 where it says "present address?"

☐ The address of the place where you live now.

☐ The address where a friend lives.

☐ The address of the place where you're applying for the job.

☐ The address of the place where you worked last.

If you have an uncle who works for this place where you would like a job and your uncle told you about coming here to apply for it, where would you write your uncle's name?

☐ Line 1

☐ Line 8

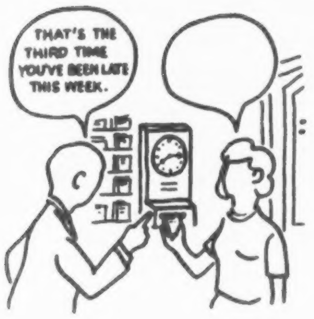
☐ Line 5

☐ Line 4

Copyright © 1967 by Educational Testing Service

Job Holding

Figure 7C


THAT'S THE THIRD TIME YOU'VE BEEN LATE THIS WEEK.

What would you say to a boss who said this to you?

☐ Sorry, I'll try to get here on time from now on.

☐ Its tough for me to make it here on time but I'll try.

☐ I do my work - what's the difference if I'm late.

Copyright © 1967 by Educational Testing Service

Figure 7D

If you feel too sick to go to work one day, call the boss and let him know.

☐ I wouldn't bother - it wouldn't make any difference to him.

☐ I might call if I happened to think of it.

☐ I would always call - otherwise he'd get sore.

Copyright © 1967 by Educational Testing Service

Motivation for Vocational Achievement (MVA)

Seventeen items, all verbal, are set up in a 5-point Likert-Scale format, as shown in the sample items below. The first 10 items consist of the item statements followed by the 5-point scale representing degree of applicability. (This format has proven successful for use with adolescents in a formal scale known as the Demos D-Scale (Demos, 1965) which is intended to identify potential school dropouts.) For the next seven items of the measure, a somewhat different format is used that was developed for assessing adolescent work attitudes in a youth-work training program (MFY, 1964). The item responses are conditional situations regarding employment and the 5-point scale represents the intensity of willingness to stay on the job. (Would you stay with it if: "The work was dirty?"; "You had to work long hours?") There are no formal measures known to exist in the specific area of vocational achievement. However, a variety of primarily informal scales that have been used for measuring "achievement orientation" are described by Kahl (1965).

Scoring for each of the two sections of MVA, with their somewhat different formats, is based upon the sum of the scaled responses, with the higher value representing the higher level of motivation. The two scores which are designated MVAI for the first 10 items and MVAII for the remaining seven, are treated separately throughout the analyses.

Figure 8A

IT PAYS TO TRY TO MAKE SOMETHING OF YOURSELF BY WORKING.

☐ Nearly always

☐ Most of the time

☐ Sometimes

☐ Very few times

☐ Almost never

Copyright © 1967 [illegible] Service

Figure 8B

Would you stay with it if:

THE BOSS WAS ALWAYS ON YOUR BACK?

☐ I would always stay with the job anyway.

☐ Most of the time I would stay with it.

☐ Sometimes I might stay with the job.

☐ I probably wouldn't stick it out.

☐ I would never stay with the job if this was the way it was.

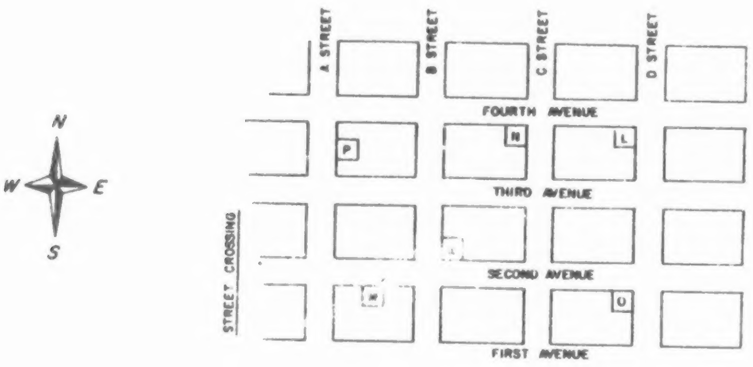
Copyright © 1967 by Educational Testing Service

Practical Reasoning--Map Reading (PRM)

This measure and the two succeeding ones probably comprise the closest approach to intellectual or cognitive measures in the battery. The rationale for their construction and use is based upon the assumption that, within a work setting, there are certain forms of skill for which an enrollee must demonstrate at least a minimum capability. These skills often center around following directions required to carry out some aspect of the job successfully. The intent here is to arrive at several indices of "reasoning" ability by utilizing material which presents problems germane to the practical work situation. It is also assumed that the practical reasoning measures represent dimensions of ability that are somewhat independent of conventional intellectual measures since they are not as highly dependent upon verbal reading skills. The 10-items in a four-choice, multiple-choice, format that constitute the Map-Reading measure, are based upon questions relating to an assumed job that requires the respondent to make deliveries in an area represented on a map (shown below with a sample item).

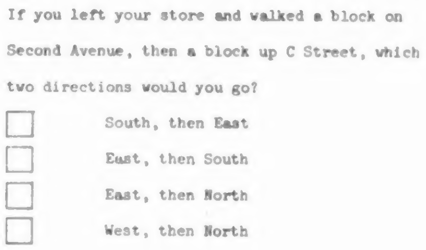
Scoring consists of the total number of items answered correctly.

Figure 9A



Copyright © 1967 by Educational Testing Service

Figure 9B



If you left your store and walked a block on Second Avenue, then a block up C Street, which two directions would you go?

☐ South, then East
☐ East, then South
☐ East, then North
☐ West, then North

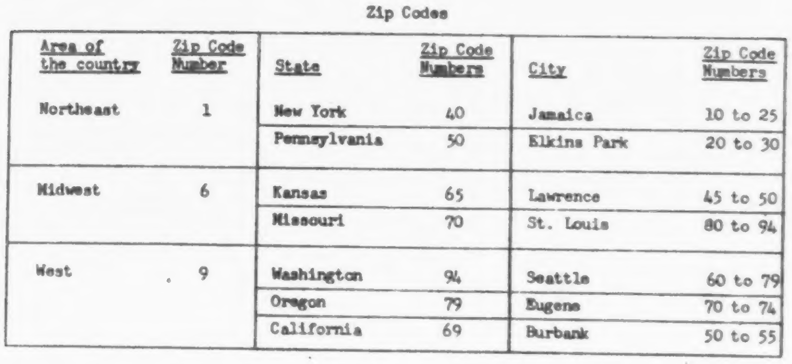
Copyright © 1967 by Educational Testing Service

Practical Reasoning--Zip Coding (PRZ)

In this 10-item measure the responses to the multiple-choice items are based upon the respondent's ability to "sort" mail (in an assumed post office job) based upon zip-code type of information. The arbitrarily designed zip-code chart and a sample item are shown below.

Scoring consists of the total number of items answered correctly.

Figure 10A

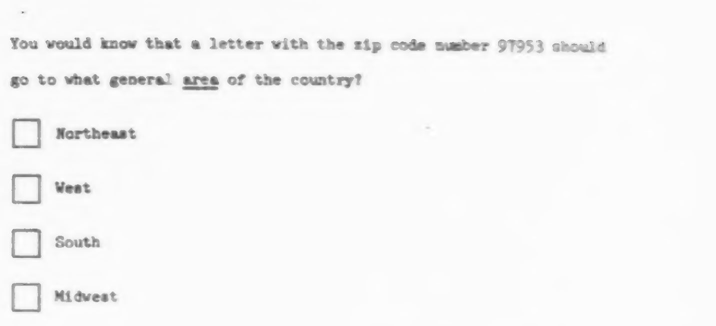


Zip Codes

Area of the country	Zip Code Number	State	Zip Code Numbers	City	Zip Code Numbers
Northeast	1	New York	40	Jamaica	10 to 25
		Pennsylvania	50	Elkins Park	20 to 30
Midwest	6	Kansas	65	Lawrence	45 to 50
		Missouri	70	St. Louis	80 to 94
West	9	Washington	94	Seattle	60 to 79
		Oregon	79	Eugene	70 to 74
		California	69	Burbank	50 to 55

Copyright © 1967 by Educational Testing Servi

Figure 10B



You would know that a letter with the zip code number 97953 should go to what general area of the country?

☐ Northeast

☐ West

☐ South

☐ Midwest

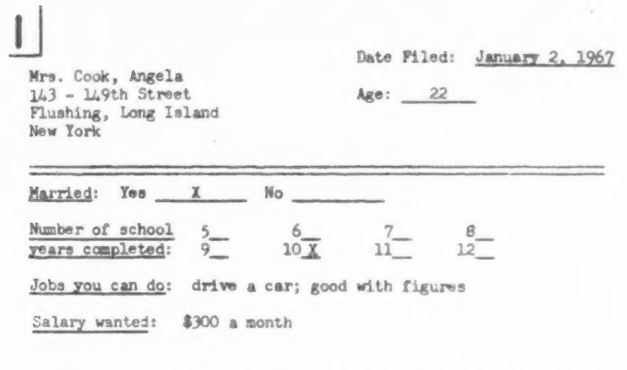
Copyright © 1967 by Educational Testing Service

Practical Reasoning--File Card Sorting (PRF)

This is also a 10-item, multiple-choice, measure. Attached to the test booklet is an envelope containing eight numbered file cards that represent information about job applicants. The respondent holds an assumed position in an employment office and is asked by his boss to sort out individual applicants based upon characteristics given on the cards, so that these men and women can be considered for various jobs as described in each item. (A sample card and item are shown below.)

Scoring consists of the total number of items answered correctly.

Figure 11A



1

Mrs. Cook, Angela
143 - 149th Street
Flushing, Long Island
New York

Date Filed: January 2, 1967

Age: 22

Married: Yes X No

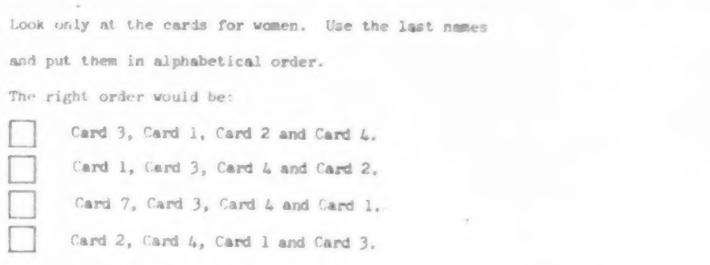
Number of school years completed: 5 6 7 8 9 10 X 11 12

Jobs you can do: drive a car; good with figures

Salary wanted: $300 a month

Copyright © 1967 by Educational Testing Service

Figure 11B



Look only at the cards for women. Use the last names and put them in alphabetical order.

The right order would be:

☐ Card 3, Card 1, Card 2 and Card 4.

☐ Card 1, Card 3, Card 4 and Card 2.

☐ Card 7, Card 3, Card 4 and Card 1.

☐ Card 2, Card 4, Card 1 and Card 3.

Copyright © 1967 by Educational Testing Service

Enrollee Rating Scale (ERS)

The 15-item, 5-point rating scale is intended as a measure of the N.Y.C. enrollee's feelings about the particular project in which he is enrolled. It covers three aspects of the N.Y.C. project having to do with the work supervisor, the counselor and the N.Y.C. project in general. The use of this measure as an evaluative tool assumes the possibility of enrollee change for what can be considered an important phase of his attitudes. The items are phrased in language understandable to the enrollees since they were collected from interviews with a sample of male and female enrollees at six projects. This material was obtained in response to a question dealing with what the enrollee liked or disliked about their counselor, the work supervisor, and the N.Y.C. project in general. Specific item statements, derivable from the interviews, were placed on cards and those 15 occurring with the greatest frequency constituted the ones chosen for the scale.

Scoring for the scales is based upon the sum of the 1 to 5 item weights ("unfavorable to favorable") over all of the 15 items.

Figure 12A

THE WORK SUPERVISOR GIVES YOU WORTHWHILE WORK, WHERE YOU CAN LEARN SOMETHING.

☐ This is never the way it is.
☐ This is usually not the way it is.
☐ Sometimes this is so.
☐ This is pretty much the way it is.
☐ This is always the way it is.

Copyright © 1967 by Educational Testing Service

Figure 12B

THE COUNSELOR TRIES TO BE HONEST AND FAIR WITH YOU.

☐ This is never the way it is.
☐ This is usually not the way it is.
☐ Sometimes this is so.
☐ This is pretty much the way it is.
☐ This is always the way it is.

Copyright © 1967 by Educational Testing Service

Counselor and Work Supervisor Rating Scales

Two scales of 11 items each serve as the criterion measures for the study--one is the Counselor Rating Scale and the other the Work Supervisor Scale. Since these are the two adults who undoubtedly have the greatest amount of personal contact with the enrollee during his tenure in N.Y.C., as well as the most relevant quality of contrast (i.e., in the counseling and job situations), they would appear to constitute the most reasonable evaluators of an enrollee's status and progress at any point in the program. All of the items in the scales are shown in Appendix C. The material for the items was collected from interviews with counselors and work supervisors at six N.Y.C. projects and the items constructed from the interviews deal with favorable and unfavorable enrollee characteristics as seen by these two groups. Counselors and work supervisors were asked to make these evaluations independently of one another, and neither group had any knowledge of enrollee performance on the battery of measures.

Scoring of each scale is based upon the sum of the 1 to 5 item weights, along the unfavorable-to-favorable continuum, over all 11 items, with a weight of "1" always assigned to the unfavorable end of the continuum and a "5" to the favorable end.

III. METHOD

A. Sample

A total of 256 Neighborhood Youth Corps enrollees in the Out-of-School program constituted the sample of respondents for the study. Of these 123 were males and 133 were females. Administration of the measures was carried out separately for male and female groups. Projects were requested to provide enrollees who had been in the program for about one month and no more than three months. Thus, the enrollees represented a sample in the earlier phases of their N.Y.C. participation. This not only served as somewhat of a control on uniformity of the group's characteristics, but was intended primarily to provide a minimum level of counselor and work supervisor acquaintance with the enrollee for rating purposes, as well as enrollee acquaintance with the program for his rating of N.Y.C.

Eleven N.Y.C. projects cooperated in the study, nine of which are located in New York, New Jersey, or Pennsylvania and can be considered as drawing their enrollees from exclusively urban areas. The other two projects are located in southern rural areas--one in Maryland and the other in Alabama.

Approximately 85% of the male sample were Negro and 15% were white or of Puerto Rican ethnic background. Of the female group, approximately 90% were Negro and 10% white or Puerto Rican. All Spanish-speaking enrollees in the sample were sufficiently bilingual to comprehend the material with a minimum of additional assistance

The sample cannot be construed in any way as representative of the national composition of the N.Y.C. Out-of-School program--not only because of geographic restriction but, more obviously, since it lacks representation of other ethnic groups that make up the enrollee population (e.g., American Indians, Mexican-Americans).* However, the sample is considered adequate for the purposes of this initial phase of development for the evaluation measures. Criterion ratings were obtained from the N.Y.C. counselors for 120 of the 123 male enrollees and 129 of the 133 female enrollees. Work Supervisor ratings were made available for 114 of the males and 121 of the females.

B. Administration

An informal tryout of the items for each measure was first conducted with a small group of three male and four female enrollees at one N.Y.C. project. This was held on a group-discussion basis and was intended to check estimated time requirements for administration, clarity of instructions and areas of ambiguous item content. Modifications in instructions and items, resulting from this pretest tryout, were incorporated in the test booklets.

The enrollee measures were administered at each N.Y.C. project to groups ranging in size from four to 15 enrollees, with male and female groups tested separately. These are untimed measures, so that administration is relatively unhurried. However, somewhat of a built-in pace is imposed by virtue of the examiner having to read every test item aloud to the group. The first seven measures were administered over a period averaging about 65 minutes, followed by a break of some 15 to 25 minutes and a second session of approximately one hour for presentation of the remaining six measures. The order of administration for the 13 measures is the same order in which they have been described in Section II of this report.

---

*Racial or ethnic characteristics are not utilized in the data analyses of this study; nor are they considered pertinent to the study purposes. It is also the preference of the project personnel that racial or ethnic information not be used in identification of individual enrollees. Thus, the proportions provided here are based upon examiner observations and only intended to add to general information on sample composition. Furthermore, individual enrollees are not identifiable by ethnic or racial characteristics in the data records.

Directions used by the examiner are shown in Appendix B. As is indicated in these directions the enrollee was made aware of the fact that the tests had no effect upon his status in the N.Y.C. program nor was he compelled to respond to any material considered, by him, to be inappropriate.

Special Problems

Certain aspects of test administration, with a group of adolescent school-dropouts who are minority group members from a poverty background, merit discussion and clarification. This is because problems can arise, that are not within the domain of those customarily encountered in the psychological group-testing situation, which apply primarily to the male enrollee groups. The female enrollees are fairly cooperative and amenable to standardized group testing procedures. With this particular population of males, who have experienced academic or disciplinary difficulties throughout much of their school careers, any group testing situation (no matter how contrived) becomes analogous to a classroom situation with the examiner in an authority position (no matter how he plays the role).

In the testing situations for this study, the primary motivation for enrollee participation was based upon their receiving the regular hourly wage over the several hours spent as test respondents, instead of working at their customary job sites. In effect, there was no more compulsion for an individual to take part in a test session than there is for him to remain enrolled in the N.Y.C. program. Thus, a number of enrollees scheduled for the test sessions did not appear; others (about six individuals) simply left during the session. Some degree of coaxing was required to prevent any large number from leaving, particularly after the mid-break period.

Examiner-enrollee interaction during a session, with its oral presentation of the tests, presents a considerably different atmosphere than encountered in testing middle-class young adults and few prior skills in test administration are applicable. There is almost constant interplay that involves "testing" of the examiner by a variety of frustrating behaviors calculated to bring about open hostility. Any show of anger in such a situation could lead to a breakdown of the test session (e.g., any attempt to reprimand an individual or telling someone to leave could lead to the entire group reacting unfavorably). The sessions are often marked by a mutually understood "put on" which is somewhat peripheral to the test process *per se*, but in which both the examiner and the enrollees remain aware of the "rules of the game."

The examiner's race may play a role in affecting the administration and subsequent results. However, it is suggested that this is less critical than the individual examiner's own mode of interacting with the group members. (The point is certainly open to further research consideration for this particular group.) Nevertheless, there remains a question of examiner contribution to such a testing situation wherein the examiner's ability to focus the respondents' attention on the test items--and generally to maintain the motivation needed for consistent response--cannot always be assumed.

An additional special problem worth recognizing is that the adequacy of a test may be dependent upon the capability of items to communicate meaningfully with some intended group; but communication is a two-way street. Although the items of the present evaluation test battery may be understandable to an N.Y.C. enrollee, they may not always be responded to in a way that reflects such comprehension (i.e., there may be a lack of willingness by the enrollee to provide consistent or meaningful responses). Whether or not this is so can only be answered, to at least some extent, in the pattern of results to be considered in the next section (Section IV) of this study.

C. Data Analyses

The approach to data analysis is, of course, tailored to the major purposes of this study. In essence, results derived from data obtained with the battery of preliminary evaluation measures are to be used to select the most acceptable test material, to suggest modifications to the existing measures and to recommend further development requirements.

The first phase of the analysis is devoted to determining item characteristics for each measure. This includes item means and variances, item intercorrelations (where such intercorrelations are useful) and the correlation of items with total-test scores.

Decisions for retaining or dropping the obviously unusable items (and/r complete measures) can be made at this level and the modified test material incorporated in the subsequent analyses. At the same time, it is possible to determine whether the rural enrollee sample (constituting 23% of the total) was sufficiently similar to the urban sample in its responses to allow the data from the two groups to be combined without serious attenuation of test validity or reliability.*

The second phase of analysis includes (a) determination of the validity of each measure against the counselor and work supervisor criterion rating scores as well as an analysis of these criterion ratings themselves, (b) estimation of reliability in the form of internal consistency--for each enrollee measure, and (c) interrelationships between the measures of the battery as well as an analysis of the underlying dimensions, or factors, that make up the enrollee measures.

Various specific analyses of the individual measures are also of interest and can allow for confirmation of prior research results (e.g., in the area of vocational choice) or provide additional findings and worthwhile research hypotheses relevant to a disadvantaged adolescent group.

Missing data occurred throughout the sample. In some few cases this occurred for complete booklets but, primarily, these failures to respond were for individual items, so that missing-data analyses had to be applied.

IV. RESULTS AND DISCUSSION

This section presents study results, first in terms of the overall findings with respect to the item data, the comparison of the rural and urban subsamples, validity and reliability of the enrollee measures and the characteristics of the criterion rating scales. The two final segments of the section are devoted to the meaning of the pattern of intercorrelations between the various measures of the battery and more detailed examination of the meaning of the results for each of the specific enrollee measures.

---

*A desire to improve the representativeness of the enrollee sample prompted inclusion of the rural group as well as some assumptions by N.Y.C. headquarters personnel about differences in test response for rural and urban groups. Thus, it is necessary to consider the pattern of responses for each group separately and the differences between them.

A. Item Analyses

Item means and variances were computed for each of the 13 enrollee measures, along with the item-total test intercorrelations for the rural and urban samples (within the male and female subgroups separately). This four-way breakdown of the item data is presented for all items in Appendix A. The intention is to consider the overall pattern of item suitability for urban and rural groups and to eliminate items which show characteristics in both groups of: (a) little or no variance, (b) overly high item means (or too easy in terms of difficulty level, for those items scored on a "right-wrong" basis) and (c) lack of positive item-total test intercorrelations indicating their failure to contribute to reliability of the overall measure.*

In general, the approach to item or test elimination was highly conservative because of the preliminary nature of the measures at this phase of study. So that only items which were particularly discrepant in terms of acceptable characteristics, for both the rural and urban samples, were eliminated from subsequent analysis. Items chosen for elimination are shown by asterisks in Appendix A. Measures in which one or more items were eliminated, and the justification, are as follows:

---

*The item-total correlations of Appendix A were not corrected for spurious overlap and should only be used as a general guide in item selection decisions, as they were in the present instance.

Job Knowledge: Four items in the male version were eliminated; three because of difficulty levels (p values) of 1.00 in either rural or urban samples (i.e., everyone answered the item correctly), and one because of negligible item-total test correlations. Of the four items, three cover questions about tools used for the jobs pictured and even the one item remaining in the "tool" category is at a p value of greater than .95 for the rural and urban samples. Obviously the entire category should be considered for elimination in future development of the measure. Five items were eliminated from the female version of the Job Knowledge test; four because of difficulty values of 1.00 and another owing to a negligible item-total correlation. These five items are not found to be primarily in any one job knowledge category.

Self-Esteem: Elimination of one item each, in the male and female versions, was made on the basis of a lack of item-total test correlation; with no particular aspect of the item content to explain the inadequacy of these two items.

Deferred Gratification: For the female version only, there was a single item eliminated because of essentially zero item-total test correlations in the rural and urban samples.

Job Seeking and Holding Skills: A decision to drop a single item, and only in the female version, was based upon its p value of approximately 1.00 in both rural and urban samples. The item is one from the first 17 items of the booklet which is the Job Seeking Skills portion of the measure, and is a question about information required for an employment application blank.

Practical Reasoning-File Cards: This is the one measure for which all items as a group could be considered unusable. The result occurred only for the female sample (urban and rural) and it is apparent that the measure represents a task far too easy for female enrollees. The expected variance restrictions (for items at an extremely easy level of performance) are also apparent from the item data presented. Thus, this measure was dropped in the subsequent analyses of results for the female sample.

Overall, relatively few items were eliminated and the measures would appear, on the basis of their patterns of item characteristics, to be applicable to a rural as well as an urban sample of N.Y.C. enrollees.

B. Urban-Rural Comparison

Prior to the major phase of the analysis, it was considered worthwhile to compare results for the rural and urban samples, based upon total mean scores for each of the enrollee measures. First, because as research findings, such differences are of interest in their own right. But, the comparison was also made because of its value for future development and utilization of the measures, as well as to determine whether or not the differences are minimal enough to justify combining the two subgroups into a single sample for data analysis purposes in the present study. Where significant differences do occur for particular measures, it is necessary to consider possible attenuation of correlations when combining the samples and the advisability of dividing rural and urban samples in future work with the measures. (Hopefully, though, there would be attempts to cross-validate any such differences with a larger rural sample than was available here.)

Table 1 presents the means and standard deviations obtained on 15 scores from the 13 enrollee test booklets. (Note that the two additional scores result from presenting separate values for the Job Seeking and Job Holding portions of the "Job Seeking and Holding Skills" booklet and a Difference Score based upon Vocational Aspirations and Vocational Plans.) Of the 30 tests of mean differences (15 each for the male and female groups) there were only two that reached a degree of significance at the .01 confidence level for the male group and three for the female group (based upon two-tailed tests). The .05 level differences are shown in Table 1 but considered at too low a confidence level to be worth stressing with the present samples. Thus, urban males can be seen to score significantly higher on the Map Reading subtest and rural males on the overall mean for Interest in Vocational Tasks items. (This latter difference indicates only a tendency of rural males to select any occupational tasks as more acceptable regardless of the particular occupational clusters in which they belong.) Rural females achieved significantly higher mean Self-Esteem scores and gave higher ratings to their N.Y.C. projects on the Enrollee Rating Scale. In summary, the pattern of item characteristics along with the relatively small number of significant mean differences that occur when urban and rural groups are compared (as well as the proportionally small rural sample size) lead to reasonable justification for combining the two samples. The remainder of the data analyses are therefore based upon combined rural and urban samples with separate analyses only for male and female subgroups.

Table 1

Enrollee Measures: Means and Standard Deviations for Rural and Urban Samples

Measure			Male		Female	
			Rural (N≅29)	Urban (N≅94)	Rural (N≅30)	Urban (N≅103)
Job Seeking Skills		Mean	9.76	9.95	10.00	11.71*
		SD	3.30	3.29	3.70	2.59
Job Holding Skills		Mean	29.55	27.86*	31.03	30.58
		SD	3.22	3.33	1.54	2.33
Enrollee Rating Scale		Mean	59.83	55.95	67.83	62.43**
		SD	11.75	12.82	6.50	9.16
Attitude toward Authority		Mean	41.41	42.41	42.85	41.31
		SD	9.02	8.34	8.96	8.15
Self-Esteem		Mean	37.07	37.76	39.73	37.99**
		SD	3.14	3.31	1.87	2.95
Deferred Gratification		Mean	53.03	51.76	53.63	53.39
		SD	7.99	7.50	6.16	5.97
Motivation for Vocational Achievement (I and II)		Mean	40.00	39.18	40.83	40.60
		SD	5.22	5.64	4.13	4.68
Job Knowledge		Mean	22.59	22.86	21.17	22.23
		SD	3.98	4.34	4.25	3.27
Vocational Interest		Mean	79.17	72.30**	64.53	62.55
		SD	11.67	11.39	8.55	11.78
Practical Reasoning	Map Reading	Mean	4.34	5.52**	4.73	4.98
		SD	1.80	2.40	1.87	2.15
	Zip Coding	Mean	5.52	6.48	7.41	7.16
		SD	2.69	2.63	2.43	2.16
	File Card Sorting	Mean	6.48	7.36	8.50	8.95
		SD	2.56	2.29	2.08	1.65
Vocational Plans (Discrepancy)		Mean	5.62	6.24	7.23	6.84
		SD	1.93	2.51	2.36	1.99
Vocational Aspirations (Discrepancy)		Mean	6.10	6.96	7.80	7.59
		SD	2.94	2.88	2.76	2.12
VA-VP (Difference)		Mean	7.10	6.05	7.90	9.31
		SD	5.79	6.40	5.48	5.90

---

*Difference between means significant at .05 level.
**Difference between means significant at .01 level.

C. Reliability and Validity of the Enrollee Measures

The internal consistency coefficients, which serve as estimates of reliability for the measures, are presented in Table 2 for the male and female samples. Several approaches to computing internal consistency are utilized. As indicated in the table, they are based upon either the Spearman-Brown formula applied to a split-half correlation; Kuder-Richardson (21) estimates for items scored on a right-wrong basis or, where there were very few items in a measure, the average item-intercorrelation was stepped-up (by the Spearman-Brown formula) to the number of items in that measure.

Reliability estimates are, for the most part, reasonably acceptable--in some instances substantial--despite the relatively small number of items comprising the various measures. The obvious exception is the overall Deferred Gratification score with reliability estimates of .37 for males and .29 for females. However, it should be recalled that the measure was constructed on the basis of four subscales of four items each. When the reliabilities of these subscales are computed (average item intercorrelation of the subscale stepped-up four times) it becomes clear from Table 2 that the attenuating effect on the total DG measure probably results from the subscale designated as "Delay of Reinforcement" which has, essentially, zero reliability. The interitem correlations for the subscales will be examined in detail in the later discussion of results for the DG measure.

Of note is that the reliability estimates for Vocational Plans and Vocational Aspirations are a measure of the intercorrelations between pairs of items at each of eight status levels. So that, in effect, these

Table 2

Reliability Estimates for Enrollee Measures

Measure	Males $(N \cong 123)$ (r_{tt})	Females $(N \cong 133)$ (r_{tt})
Job Knowledge (JK)[b]	.75	.72
Vocational Plans (VP)[a]	.81	.57
Vocational Aspirations (VA)[a]	.78	.67
Vocational Interest (VI)		
(1) Clerical[c]	.62	.80
(2) Service[c]	.69	.65
(3) Technical[c]	.39	.71
(4) Outdoor[c]	.65	.78
(5) Science[c]	.60	.69
(6) Business[c]	.56	.64
(7) Aesthetic[c]	.67	.45
Attitude toward Authority (AA)[a]	.64	.87
Self-Esteem (SE)[a]	.50	.60
Deferred Gratification (DG)[a]	.37	.29
(1) Freedom of Spending[c]	.74	.69
(2) Hasty Aggression[c]	.64	.56
(3) Affiliation[c]	.61	.61
(4) Delay of Reinforcement[c]	.07	.06
Job Seeking Skills[b]	.65	.66
Job Holding Skills[a]	.74	.72
Motivation for Vocational Achievement (MVAI)[a]	.81	.58
(MVAII)[a]	.81	.64
Practical Reasoning		
(1) Map Reading (PRM)[b]	.60	.47
(2) Zip Coding (PRZ)[b]	.75	.66
(3) File Card Sorting (PRF)[b]	.71	--[d]
Enrollee Rating Scale[a]	.96	.90

---

[a] Split Half; stepped up by Spearman-Brown formula.
[b] Kuder-Richardson (21) estimates.
[c] Average Interitem $\overline{r}$ stepped up by Spearman-Brown formula.
[d] Dropped from use with female sample.

are the closest of all the coefficients to an alternate (or parallel) form type of reliability estimate. They also represent a measure of the degree to which consistency of status underlies the vocational choices. This degree of consistency of choice can be seen in Table 2 to be rather similar for the plans and aspirations measures. The exceptionally high level of reliability for the Enrollee Rating Scale (r = .96 for males and .90 for females) is also worth noting and is indicative, not only of extreme consistency in judgment, but of considerable generalization or "halo" effect over the items that make up the scale. That is, if the enrollee's attitude toward the N.Y.C. project is favorable, it tends to be a highly generalized attitude over all phases of the program.

The correlations that represent validity coefficients are presented in Table 3 for the male and female enrollee samples. These relationships between the enrollee measures and the counselor and work supervisor criterion ratings can be considered as no more than low-to-moderate in size. But is is also the pattern of these correlations which is of interest for defining the measures of the battery that might be more predictive of longer-term enrollee vocational or social adjustment. They also provide clues as to the bases on which counselors and work supervisors evaluated the enrollees.

The pattern of validities is found to be fairly similar for counselors and work supervisors, although certain differences between them are worth noting--particularly in their evaluation of male and female enrollees. For both criterion raters, a clearly significant set of relationships exists between their ratings of male enrollees and performance on the three Practical

Table 3

Validity Coefficients: Correlation of Counselor and Work

Supervisor Criterion Ratings with Enrollee Measures

	Male		Female	
Enrollee Measure	Counselor (N≅112)	Work Sup. (N≅102)	Counselor (N≅129)	Work Sup. (N≅128)
Job Knowledge	.22*	.24*	.19*	.03
Att. toward Auth.	.12	.22*	.06	.06
Self-Esteem	.04	.01	.15	.21*
Def. Gratification	.27**	.17	.18	-.03
Job Seeking Skills	.32**	.35**	.26**	. .22*
Job Holding Skills	.22*	.19	.26**	.17
Mot. for Voc. Ach. I	.26**	.21*	.21*	.13
Mot. for Voc. Ach. II	.02	.03	.10	.13
Map Reading	.24*	.28**	.06	.06
Zip Coding	.20*	.37**	.25**	.08
File Card Sorting	.23*	.31**	-	-
Enrollee Rating Scale	.22*	.21*	.20*	.15
VA (Discrepancy)	-.05	-.03	-.03	-.13
VP (Discrepancy)	-.04	-.06	-.11	-.13
VA-VP (Discrepancy)	-.07	-.18	.00	-.07
Vocational Interest				
Clerical	.17	.16	.10	.03
Service	.04	-.01	.25**	-.01
Technical	-.11	-.02	.00	-.01
Outdoor	.04	.01	.08	.01
Science	.09	.09	.15	.00
Business	.07	.14	-.04	-.09
Aesthetic	.04	-.12	.07	-.19

---

*r significantly greater than zero at .05 level.
**r significantly greater than zero at .01 level.

Reasoning Subtests; with larger correlations for the work supervisor ratings. Apparently, the male enrollees who have the ability to follow directions in carrying out tasks that also have some reading-skills component (as is characteristic of the Practical Reasoning measures) are seen as superior. For females, however, the abilities reflected in this type of measure are not significantly indicative of performance ratings by work supervisors. Only the direction-following and numerical sort of task that makes up the PR Zip Code subtest bears a low positive relation to counselor ratings of the females ($r = .25$).

Consonant with these findings are the significant correlations between rater evaluations and the Job Seeking Skills subtest (especially for males) with an r of .32 for counselor ratings and .35 for work supervisor ratings. These 17 items, it should be recalled, consist primarily of employment want-ads and sections of an application blank. They are also the most heavily weighted of all the enrollee measures in a reading requirement and constitute virtually the only material in the battery which the enrollee must read entirely on his own. Thus, an individual who is superior in the skills required for this set of items tends, uniformly, to be judged as a better enrollee by counselors and work supervisors. The one other measure that requires what might be considered some component of intellectual capability (although not specifically verbal to any great extent) is the Job Knowledge measure. This also appears as positively related to counselor and work supervisor ratings for males but with lesser validity for the females.

Of the attitudinal types of measures, the best validity is generally found with the counselor rating criterion. Among males, the significant correlations, in order of magnitude, are found for Deferred Gratification (r = .27), Motivation for Vocational Achievement: I (r = .26), Job Holding Skills (r = .22), and the Enrollee Rating Scale (r = .22). The pattern of significant relationships closely parallels those found for counselor ratings of the female enrollees. Work supervisor ratings of males present a pattern of validities reasonably similar to that of counselors (i.e., significant, but low, correlations with JH, MVA:I and ERS--although not quite significantly related to DG). One difference worth noting is that Attitude toward Authority reaches a significant level with the work supervisor criterion ratings of males (r = .22). It would seem logical that work supervisors place greater stress on this type of conformity among the male enrollees.

From Table 3 it can be seen that with one exception, the four-item job clusters of the Vocational Interest Measure achieve no particular validity for either males or females. The exception is that females who tend to score high on the Service cluster also tend to be rated higher by counselors (r = .25).

The next section provides a more comprehensive view of the pattern of validity and its meaning for the development and use of the enrollee measures.

D. Intercorrelations and the Factor Pattern

The matrices of intercorrelations for all of the evaluation measures in the enrollee battery are shown in Tables 4 and 5. The 22 x 22 matrix for the male and the 21 x 21 matrix* for the female enrollees were each factor analyzed in order to determine the underlying factors that make up the test scores. From the pattern of interrelationships and the clustering of the various measures, a more precise statistical summary of the dimensions tapped by each measure can be identified, as well as pertinent clues to the most effective measures utilized in this preliminary development phase.

For both factor analyses, seven factors were extracted by a principal components solution and the factors were then rotated using a varimax rotational method (Kaiser, 1958).** The two criterion rating scales were added to the rotated matrix by factor extension in order to provide an overview of the validity contribution of each dimension.

In the male enrollee sample, which will be considered first, four of the factors had a sufficient number of measures with loadings at a reasonable enough level to allow for interpretation. Three of the four are rather clearly designed as Intellectual Skills (Factor I), General Vocational Interest (Factor II) and Positive Attitudes (Factor III). The fourth factor, which is defined by relatively few of the measures in the battery, is probably best conceived of as one indicating a Negative Self Image (Factor IV).

---

*The Practical Reasoning File Card subtest was eliminated for females as indicated in Section IV A.

**The seven factors accounted for 69% of the total variance in the matrix for the male sample and 66% of the total for the female sample.

Table 4

Total Score Correlation Matrix[a]

Enrollee Measures (Males N=123)

	Job Knowledge (JK)	Att. toward Auth. (AA)	Self-Esteem (SE)	Deferred Grat. (DG)	Job Seeking Skills (JS)	Job Holding Skills (JH)	Mot. for Voc. Ach. I (MVAI)	Mot. for Voc. Ach. II (MVAII)	Map Reading (PRM)	Zip Coding (PRZ)	File Card Sorting (PRF)	Enrollee Rating Scale (ERS)	Int. in Voc. Tasks (VI)							Vocational Aspir. (Status Discrep.)	Vocational Plans (Status Discrep.)	VA-VP (Dif.)
													Clerical	Service	Technical	Outdoor	Science	Business	Aesthetic			
Job Knowledge (JK)		13	19	28	58	49	46	15	45	45	46	36	13	-15	03	-26	19	12	-10	-06	-05	06
Attitude toward Authority (AA)			22	30	37	40	16	-12	31	20	25	24	-09	-09	10	-04	04	10	01	-08	-08	06
Self-Esteem (SE)				33	17	40	22	03	03	10	10	31	01	-01	11	-16	04	04	-10	-12	-12	25
Deferred Gratification (DG)					22	29	38	16	31	30	31	24	01	-01	09	-02	13	13	01	-10	-39	13
Job Seeking Skills (JS)						38	39	-00	55	60	65	29	20	-05	07	-33	28	32	-01	-25	-05	-06
Job Holding Skills (JH)							49	28	29	30	29	34	15	02	12	-08	16	17	08	-10	-08	11
Mot. for Voc. Ach. I (MVAI)								39	30	41	38	43	16	07	12	-15	18	19	03	-22	-11	18
Mot. for Voc. Ach. II (MVAII)									10	08	06	15	11	09	13	17	04	09	11	-23	-07	-01
Map Reading (PRM)										62	62	10	09	-18	09	-23	16	10	-09	02	-05	-06
Zip Coding (PRZ)											69	29	16	-02	09	-29	24	19	02	-03	-02	07
File Card Sorting (PRF)												26	17	-03	03	-27	19	24	02	-14	-12	-04
Enrollee Rating Scale (ERS)													10	19	23	13	22	29	16	-18	00	16
Clerical														42	33	-01	52	52	40	05	21	12
Service															21	26	46	47	57	04	21	16
Technical																30	51	41	37	-04	07	19
Outdoor																	12	02	19	-11	01	-10
Science																		56	43	-01	11	13
Business																			47	-17	01	26
Aesthetic																				-07	07	19
Vocational Aspirations																					36	02
Vocational Plans																						-05
VA-VP (Difference)																						

---

[a] Decimal points omitted.

Table 5

Total Score Correlation Matrix[a]

Enrollee Measures (Females N=133)

	Job Knowledge (JK)	Att. toward Auth. (AA)	Self-Esteem (SE)	Deferred Grat. (DG)	Job Seeking Skills (JS)	Job Holding Skills (JH)	Mot. for Voc. Ach. I (MVAI)	Mot. for Voc. Ach. II (MVAII)	Map Reading (PRM)	Zip Coding (PRZ)	Enrollee Rating Scale (ERS)	Int. in Voc. Tasks (VI)							Vocational Aspir. (Status Discrep.)	Vocational Plans (Status Discrep.)	VA-VP (Dif.)
												Clerical	Service	Technical	Outdoor	Science	Business	Aesthetic			
Job Knowledge (JK)		11	10	35	58	29	25	22	35	56	06	14	15	-22	-15	04	-04	-05	05	-02	15
Attitude toward Authority (AA)			12	36	13	30	24	06	21	21	07	04	-03	-17	-10	-13	-02	-08	07	-06	05
Self-Esteem (SE)				07	13	40	31	18	20	20	28	23	11	-01	-03	15	21	03	-17	04	-06
Deferred Gratification (DG)					33	33	32	11	13	28	01	23	12	-09	-11	20	05	-03	-01	-06	04
Job Seeking Skills (JS)						36	31	25	41	61	-03	27	09	-21	-20	04	-09	-07	-07	04	05
Job Holding Skills (JH)							44	18	18	33	26	20	14	-13	-14	02	03	-02	-10	-05	19
Mot. for Voc. Ach. I (MVAI)								27	26	34	26	48	32	10	-10	19	29	12	-17	-03	05
Mot. for Voc. Ach. II (MVAII)									15	18	21	27	13	13	15	32	07	09	-07	-02	11
Map Reading (PRM)										41	09	13	17	-10	-15	01	00	-01	-05	-03	08
Zip Coding (PRZ)											06	24	13	-22	-20	-00	-05	-12	13	07	-00
Enrollee Rating Scale (ERS)												13	07	05	10	14	15	-07	-01	00	-10
Clerical													25	11	-07	37	54	20	-14	11	07
Service														43	19	39	28	47	-08	01	02
Technical															59	44	45	57	-18	-14	07
Outdoor																28	18	27	-08	-17	-03
Science																	41	38	-15	11	18
Business																		41	-22	17	-01
Aesthetic																			-22	-10	13
Vocational Aspirations																				29	-22
Vocational Plans																					-10
VA-VP (Difference)																					

---

[a] Decimal points omitted.

The pattern of loadings for Factor I, as shown below, is readily definable from the appearance of its highest loadings on the three Practical Reasoning measures as well as on the Job Knowledge and Job Seeking Skills Measures. (Loadings of about .25 or greater were considered for interpretation throughout.)

Factor I

Intellectual Skills

Measure	Factor Loading
Practical Reasoning	
File Cards	.83
Map Reading	.82
Zip Code	.80
Job Seeking Skills	.79
Job Knowledge	.62
Motivation for Vocational Achievement (I)	.39
VI: Outdoor Cluster	-.39
Job Holding Skills	.39
Deferred Gratification	.29
Attitudes toward Authority	.29
Work Supervisor Rating	.39
Counselor Ratings	.28

A number of the attitudinal measures also make a contribution (although a relatively minor one) to this factor. Their loadings obviously confirm what appears in the intercorrelation matrix, where these attitudinal measures are seen to have low to moderate correlations with the intellectual skills type of measures.

The one loading that appears for the Vocational Interest job clusters is on the "Outdoor" category. Its negative loading seems logical as an indication that male enrollees who score high on the Intellectual Skills factor tend to score low on (i.e., reject) the sort of job tasks that make up the Outdoor cluster--the tasks pictured being unquestionably lower level, or menial.

Probably of major interest in this pattern of loadings is that the criterion rating scales for counselors and work supervisors also achieve loadings of interpretable magnitude on the factor. As will be apparent from the remainder of the analysis, this is the only factor for males to which these criteria make a significant contribution. It is clear that these raters tend to be more favorable in their evaluation of male enrollees who score higher on the sort of measures that define the Intellectual Skills factor and, in addition, that it is the work supervisor who seems to place greater emphasis on this area of capability.

For Factor II (General Vocational Interest) the pattern of loadings is an obvious reflection of the intercorrelations between the job clusters of the "Interest in Vocational Tasks" measure. Favorable choices or acceptance of tasks in job clusters tend to generalize over virtually all of the VI clusters. As might be expected, the Outdoor cluster constitutes the exception and is not found to be loaded on this factor.

Factor II

General Vocational Interest

Measure	Factor Loading
VI Business	.78
Science	.76
Aesthetic	.75
Service	.73
Clerical	.70
Technical	.60
VA-VP (Difference)	.31
Enrollee Rating Scale	.24

An unusual addition to the loadings is the VA-VP Difference score which indicates that the enrollee with generally positive occupational task interests is one who has a larger difference between his vocational aspirations and plans. (For these data, as will be shown in later discussion, it also means, specifically, that the tendency is for such a difference to be in the direction of higher aspirations.) Individuals who score higher on the Enrollee Rating Scale also tend to be those who have a pattern of broader vocational interests.

Factor III has its lead loadings on the unmistakably attitudinal measures of the battery and represents a dimension of Positive Attitudes.

Factor III

Positive Attitudes

Measure	Factor Loading
Job Holding Skills	.70
Self-Esteem	.70
Enrollee Rating Scale	.66
Attitude toward Authority	.52
Motivation for Vocational Achievement (I)	.47
Job Knowledge	.39
Job Seeking Skills	.28
VA-VP (Difference)	.28

Two of the measures that are concerned primarily with intellectual skills (Job Seeking Skills and Job Knowledge) have a minor role in the factor--again, indicating the modest overlap between attitudes and intellectual skills for this group. The VA-VP Difference score also makes a minor contribution to the factor, indicating that a high score on the attitudinal dimension is accompanied by a tendency to show a larger difference between vocational aspirations and plans. It should also be noted that there are no loadings of reasonable magnitude found on this factor for counselor and work supervisor ratings. Those individual attitudinal measures that do correlate significantly with the criterion ratings apparently do so largely because of the intellectual component that they also possess.

The fourth factor found for male enrollees is defined by only five of the measures but seems reasonably interpretable.

Factor IV

Negative Self Image

Measure	Factor Loading
VI: Outdoor Cluster	.73
VA-VP (Difference)	-.53
VI: Technical Cluster	.33
Attitude toward Authority	.30
Self-Esteem	-.28

The description is essentially one of an individual who is strongly favorable in his choice of the low-level, menial type of tasks that define the Outdoor Cluster of the VI measure and who shows little difference between his vocational aspirations and plans. He also tends, to a minor degree, to favor the tasks of the Technical job cluster; to have a higher score on the Attitude toward Authority measure and a lower Self-Esteem score. In essence, this seems to define the individual with a somewhat diminished self-concept or Negative Self Image.

The factor pattern changes somewhat for the female enrollee sample, with five factors having a sufficient number of loadings of sufficient magnitude to warrant their interpretation.

Factor I is readily designated as the Intellectual Skills Factor and is certainly comparable to the one found for the males.

Factor I

Intellectual Skills

Measure	Factor Loading
Job Seeking Skills	.80
PR Zip Code	.79
Job Knowledge	.76
PR Map Reading	.67
Motivation for Vocational Achievement (I)	.33
Job Holding Skills	.29
VI Service	.27
VI Outdoor	-.26

Here again, there is the minor contribution of measures that are primarily attitudinal in content and also the negative loading for the low-level tasks of the VI Outdoor cluster on this dimension that reflects a form of intellectual skill. A unique addition to the factor for females is the VI Service cluster. Counselor and work supervisor ratings make no contribution to this factor worth considering.

Factor II corresponds very closely to the General Vocational Interest dimension found for the male sample and is designated in the same way.

Factor II

General Vocational Interest

Measure	Factor Loading
VI Technical	.83
Aesthetic	.76
Service	.71
Outdoor	.59
Business	.51
Science	.50

This generally favorable attitude toward job tasks extends over all of the occupational clusters, except for clerical, which seems to occupy a unique role for females. The girl who scores high on the clerical cluster does not tend to favor all of the job tasks indiscriminately. In fact, the relatively high positive loading on the Outdoor (menial tasks) cluster would appear to characterize this factor as one of low-level, indiscriminate task interest. Support for this contention is found to some extent in the presence of a third factor which displays its positive loadings on the VI clusters of Clerical, Business and Science and is negatively loaded on the Outdoor cluster. The factor is designated as High Level Occupational Interest and Motivation (Factor III).

Factor III

High Level Occupational Interest and Motivation

Measure	Factor Loading
VI Clerical	.69
Vocational Plans (Discrepancy)	.65
VI Business	.60
Science	.40
Outdoor	-.34
Motivation for Vocational Achievement (I)	.30

The female enrollee defined here is one who scores high in interest on the vocational tasks in clerical, business and science areas, is unfavorable toward the outdoor tasks and scores high on Motivation for Vocational Achievement. However, the high positive loading on the Vocational Plans discrepancy score indicates a somewhat puzzling tendency toward inconsistency in status choice.

A factor of particular importance for the female sample is one that is easily interpretable as Positive Attitudes (Factor IV). The added importance stems from its being the only factor that begins to show a loading of any reasonable level on a criterion rating scale.

Factor IV

Positive Attitudes

Measure	Factor Loading
Enrollee Rating Scale	.73
Self-Esteem	.72
Job Holding Skills	.44
Motivation for Vocational Achievement (I)	.44
VA-VP (Difference)	-.27
Motivation for Vocational Achievement (II)	.26
———	
Work Supervisor Rating	.27

This highly positive attitude toward the N.Y.C. program (Enrollee Rating Scale) and toward oneself (Self-Esteem), along with favorable attitude toward the job situation (Job Holding Skills) and toward employment in general (Motivation for Vocational Achievement), seems to define the female enrollee who is more likely to be seen favorably by the work supervisor. This is in notable contrast to the factor pattern for males, for whom favorable evaluations tend to be related to measures that define the Intellectual Skills factor with no such validity appearing for the attitudinal factor. (The female sample also seems to produce a "cleaner" distinction between the attitudinal and intellectual components--so that attitudes do not appear as highly interrelated with intellectual skill as for males.)

The fifth and final factor interpreted here is best considered as one that defines Social Conformity (Factor V).

Factor V

Social Conformity

Measure	Factor Loading
Attitude toward Authority	.78
Deferred Gratification	.75
Job Holding Skill	.48
Motivation for Vocational Achievement (I)	.39

Although significantly loaded on only four measures, its interpretation seems clear. More favorable attitudes toward authority and toward the social values that define gratification deferment, constitute the primary bases for defining the factor. Logically supplementing these are the positive tendencies toward social conformity that would seem to mark higher scores on the Job Holding Skills and Motivation for Vocational Achievement measures.

E. The Criterion Scales

Since these two rating measures serve as the only independent standards (i.e., criteria) of enrollee capability, it is desirable to consider their scale characteristics in some detail as well as certain aspects of their value as criterion measures.

The 11 items that make up each scale (Appendix C) can be seen to consist of 10 fairly specific forms of enrollee behavior and one final item representing a summary evaluation of enrollee capability. The means, standard deviations and item-total score correlations are presented in Tables 6 and 7.

Table 6

Counselor Rating Scale: Item Characteristics

	Male (N=112)			Female (N=129)		
Item	Mean	S.D.	Item-Total r	Mean	S.D	Item-Total r
1)	3.91	.88	.52	3.92	.83	.78
2)	3.39	1.28	.54	4.17	.99	.70
3)	4.14	1.12	.63	4.18	.95	.55
4)	3.97	.99	.70	3.76	1.15	.77
5)	3.19	1.14	.68	3.61	1.21	.83
6)	3.40	1.09	.27	4.37	.75	.67
7)	3.54	1.01	.62	3.76	1.21	.45
8)	3.82	1.15	.80	3.64	1.14	.81
9)	3.65	1.23	.46	4.37	.98	.57
10)	3.45	1.37	.70	3.97	1.13	.66
11)	3.54	1.14	.85	3.93	1.10	.85

Table 7

Work Supervisor Rating Scale: Item Characteristics

	Male (N=102)			Female (N=128)		
Item	Mean	S.D.	Item-Total r	Mean	S.D.	Item-Total r
1)	4.22	.90	.59	4.24	.77	.84
2)	3.50	1.23	.50	4.13	1.16	.57
3)	4.26	1.03	.64	4.36	.79	.46
4)	4.21	.94	.71	4.16	.80	.78
5)	3.45	1.28	.80	3.78	1.20	.76
6)	3.75	1.06	.55	4.55	.68	.67
7)	3.72	1.12	.47	3.96	1.25	.55
8)	4.07	1.06	.81	3.81	1.04	.79
9)	3.97	1.14	.68	4.60	.89	.45
10)	3.87	1.28	.63	4.26	.96	.71
11)	3.94	1.16	.86	4.17	.96	.87

Item-total score correlations are substantial and reflect, to some extent, the extremely high split-half reliability estimates found. These estimates (stepped up by the Spearman-Brown formula) are .94 for counselor ratings and .98 for supervisor ratings. Even though such split-half reliabilities tend to overestimate somewhat (Thorndike, 1951), they nevertheless indicate a gross "halo" or generalization effect on the part of the raters. Thus, a favorable attitude toward an enrollee tends to result in a favorable rating no matter what the specific behavior being evaluated. The conclusion is further supported by the item-total scale intercorrelations for the summary item in the scale (item #11). Correlations of .85 for both male and female samples on counselor ratings and .86 and .87 for these samples on the work supervisor scales indicate that the one summary item is almost as effective a measure as the remaining 10 items combined. Not only is there gross generalization in the evaluations, but the fairly extreme leniency shown by the means of Tables 6 and 7 also point to a difficulty in obtaining discriminating judgments for the various characteristics rated--especially on the part of work supervisors.

Counselor mean rating scores, overall, are 3.6 for males and 3.9 for females on the 5-point scaled items. Work supervisor mean ratings are 3.9 for males and 4.1 for females. This rating leniency also tends to restrict variability for the scale items. One could assume that the work supervisors simply see enrollees as distinctly superior in all characteristics covered by the items. But this seems gratuitous and it is more likely that they are reluctant to rate enrollees unfavorably on a formal measure. More careful techniques of administering the scale could prove effective--such as instructions which stress that raters try to assign a wider range of weights for the items; plus specific assurances that the results would not affect an enrollee's status in the program. Unfortunately, there was total dependence upon N.Y.C. project counselors leaving the rating scales with the work supervisors and having them returned when completed, so that only the printed instructions on each scale were available to the work supervisor and these are minimal (Appendix C).

Agreement between counselor and work supervisor ratings of enrollees (interrater reliability) is found to be relatively modest with a correlation between the total scale scores of .36 for the males and .36 for the female sample. This rather mediocre consensus regarding an enrollee's degree of "successful" or "unsuccessful" functioning represents an argument not only for improving the rating instruments (and perhaps the judgmental ability of at least one of the groups of raters) but validation of counselor and work supervisor judgment against longer range (follow-up) performance measures for N.Y.C. enrollees.

F. Discussion of Specific Enrollee Measures

On the basis of the results already discussed, and additional data analyses, it is of value to place these findings in general perspective as they apply to each of the enrollee measures. For ease of discussion, the results are considered first for those measures found to define the intellectual skills dimension, followed by those most readily classified as attitudinal.

Job Knowledge: The measure can be considered one of the most effective of the battery in terms of its reliability (in the .70's) and its correlations with the other enrollee measures, as well as with the criterion ratings. Its relationships cut across both the other intellectual skills types of measures and the attitudinal ones. The levels of these correlations are relatively high, particularly with the Job Seeking Skills measure and the Practical Reasoning measures. For the male enrollee sample it tends to show more significant correlations with the attitudinal measures (than it does for the females) and is also more highly correlated with the criterion ratings given by counselors and work supervisors.

An additional finding that may be of value regarding enrollee knowledge of job requirements can be seen in the scores achieved on the various categories that constituted the scale--i.e., Place (where the job is performed), Nature of the Work, Working Hours, Education Required and Starting Salary (for full-time employment). The "Tools" category is not considered here because of almost no errors having been committed in the responses for that category.

The proportion of correct responses in each of the five Job Knowledge categories is as follows:

Category	Males	Females
Place	.90	.95
Nature of Work	.85	.88
Hours	.76	.78
Education	.59	.78
Salary	.51	.60

These proportions appear to speak for themselves and generally indicate that the weakest areas of job knowledge occur in the categories of Hours, Education and Salary requirements. However, such a generalization should be tempered by noting that, for the fairly common occupations depicted in the items of this measure, it is not easy to construct reasonable distractors for items that ask "Where would you work on this job?" On the other hand, item difficulty level is much more readily controlled when the distractors represent salary range. In any event, the possibility of eliminating the "Place" category of items would be open for consideration on the basis of these data.

Practical Reasoning Subtests: The three measures that have been intended to assess practical reasoning provide a somewhat different picture of results for the male and female groups. For males the three measures (Map Reading, Zip Code and File Card Sorting) form a highly interrelated cluster and would appear to measure skills with considerable overlap. Their level of reliability can be considered fairly good and validity was shown to be at a consistently significant level for both counselor and work supervisor criterion ratings--with appreciably better validity for their relationships with work supervisor ratings.

As measures that show the highest loadings on one of the major factors of the battery they serve to define a clear intellectual skills factor, although they are also found to be moderately related to certain of the vocationally-oriented attitudinal measures (i.e., Job Holding Skills and one of the Motivation for Vocational Achievement subtests). There is reasonable evidence for a favorable potential role, by these Practical Reasoning measures, as evaluative instruments and as predictors of male enrollee performance in a job setting.

Among female enrollees, the Practical Reasoning measures did not prove as successful, first, because of the complete inadequacy of the File

Card Sorting subtest as a measure (far too easy a task) and, in addition, because the two remaining PR subtests (Map Reading, Zip Coding) show a less dominant effect in defining an intellectual skills factor. They are also not as consistently related to the criterion ratings, as in the male sample, and they are not as highly correlated with one another ($r = .41$). Nevertheless, the item characteristics of both measures are suitable and their pattern of significant intercorrelations with the other measures is extensive enough to warrant retaining both of them in the battery for future use and analyses with female enrollee samples. The more useful measure of the two for females would, generally, seem to be the Zip Code subtest.

Job Seeking Skills: This major contributor to the intellectual ability dimension of the battery appears to achieve its status primarily because of the requirement for verbal reading skills that it imposes on the respondent. At the same time, however, the Job Seeking measure possesses one of the most consistent patterns of positive relationships with the attitudinal measures. Specifically, this occurs with the Attitude toward Authority measure for males with the two attitude measures having the highest vocationally oriented content--i.e., Job Holding Skills and Motivation for Vocational Achievement (r's in the .30's) for both males and females. The reliability of the measure appears more than adequate for future use, with its r_{tt} of .65 for males and .66 for females. Of particular note, however, is that this Job Seeking Skills subtest constitutes the best overall single predictor in the battery of both the counselor and work supervisor criterion ratings.

Attitude toward Authority: A favorable pattern of item characteristics (means, variances, item-total test r's) and substantial reliability estimates, for both male and female samples, characterize this attitudinal measure. For the male enrollee sample, it shows a number of significant relationships to other attitude dimensions throughout the battery, as well as the highest single correlation of any attitude measure with the work supervisor ratings ($r = .22$). For females, although the measure does not display any particular validity, it does constitute the major influence on the fairly distinctive dimension of Social Conformity. In general, it possesses sufficient consistency, both internally and in the pattern of interrelationships with the other measures, to warrant its retention--especially for the male enrollee group where it appears to have promise for predicting later vocational adjustment.

Self-Esteem: One of the features of this measure is the relatively "pure" attitudinal aspect of its contribution to the battery. Among the males it can be seen to correlate significantly with only attitudinal scales (DG, JH, MVA: I, ERS and the VA-VP Difference score). Since measures on the attitude dimension were shown to possess little validity for males, this is found to be especially true for the Self-Esteem measure which is exclusively attitudinal in its factor contribution. As might be expected from the factor pattern for females, and from the modest validity achieved for the Positive Attitude dimension (on which Self-Esteem had the largest factor loading), it does make a strong contribution to what little validity the battery achieves for the female sample. Within the battery its intercorrelations are highest for the most vocationally oriented in content of the attitude measures--i.e., Job Holding Skills and Motivation for Vocational Achievement: I. Self-Esteem, for the female sample, also shows relatively rare significant intercorrelations with job clusters of the "Interest in Vocational Tasks" measure (r of .23 with the Clerical cluster and .21 with the Business cluster). In general, the Self-Esteem scale possesses a respectable pattern of item characteristics and a suitable level of reliability, although it might be expected to have more potential value with a female enrollee sample.

Enrollee Rating Scale: Similar in a number of its relationships to the Self-Esteem measure, the Enrollee Rating Scale was found to serve almost exclusively as an attitudinal measure. The major difference is that it results in a greater number of significant correlations with the criterion ratings. Among the measures used with male enrollees, ERS was found to have a considerable number of significant and logical intercorrelations, but is particularly weak in this respect for the female sample. Of relevance in the pattern of intercorrelations for males is that the relationships are strongest with the vocational-content type of measures--i.e., Job Knowledge (r = .36), Job Holding Skills (r = .34) and Motivation for Vocational Achievement: I (r = .42).

To some extent, the failure of the Enrollee Rating Scale to correlate with numerous other measures used for the female sample is likely to have resulted from the unusually high item mean values (i.e., rater "leniency") and consequent restriction in variance that resulted. Even with this limitation, the scale has a valuable function in the battery and would prove even more useful for females if the scale format and instructions are revised to elicit more discriminating (i.e., critical) levels of response.

Deferred Gratification: Despite the built-in multidimensionality of this measure (i.e., four separate subscales), the DG total score seems to play a logical and consistent role in its relation to the other measures of the battery. For both male and female enrollees, it is one of the attitudinal measures that displays a minor intellectual-skills component (as seen in its factor loadings) and among females it also provides a major contribution to the attitudinal dimension of Social Conformity. The level of validity for DG is found to be significant only for males and primarily in relation to the counselor criterion ratings.

Although its reliability estimates were found to be the lowest in the battery, this poor reliability for the total DG scores was shown to be attributable, for the most part, to a lack of any reasonable level of internal consistency in one of the subscales. This attenuating effect is the likely result of having incorporated the group of items that make up the subscale defined as "Delay of Reinforcement" (item numbers 4, 11, 14 and 16 in Appendix A for the DG measure). These individual items also show less than outstanding item-total test correlations in the DG measure.

The item intercorrelations for scores on each of the four subscales that constitute the DG measure are shown for the male enrollee sample. (The pattern for females is essentially the same.) The intention is to provide an illustration of the value and coherence of each of the subscales and to demonstrate the probable attenuating effect of these items on correlations computed with the total scores. Table 8 presents these item intercorrelations for the four subscales which are designated as "Affiliation," "Freedom of Spending," "Hasty Aggression" and "Delay of Reinforcement."

Cluster II designated as "Freedom of Spending" is the most coherent of the four with the highest item intercorrelations. Obviously, the "Delay of Reinforcement" cluster (IV) does not represent a subscale at all.

Table 8

Item Intercorrelations for Subscales of the Deferred Gratification Measure (Male Enrollee Sample)

N = 123

I. Affiliation

Item #

	1	9	13	15
1		.19	.22	.23
9			.43	.24
13				.38

II. Freedom of Spendin

Item #

	2	5	7	8
2		.35	.41	.22
5			.53	.41
7				.47

III. Hasty Aggression

Item #

	3	6	10	12
3		.14	.26	.12
6			.41	.22
10				.47

IV. Delay of Reinforcement

Item #

	4	11	14	16
4		-.15	-.03	.21
11			-.01	-.16
14				.02

The overall performance of the DG measure in the battery of enrollee measures (despite the probable weaknesses imposed by one group of items) indicates that this attitudinal scale is worth further development effort. This should be based upon incorporating the three best subscales with some judicious modification, or item reconstruction, in order to improve their coherence where possible.

Job Holding Skills: This second portion of the Job Seeking and Holding Skills test booklet is most aptly considered an attitudinal type of measure involving knowledge of proper behavior on a job, rather than one requiring intellectual or factual knowledge. This conclusion is verified by the relatively high relationships between JH and the other attitudinal measures of the battery. But it is also seen to have modest correlations with several of the intellectual skills measures and a corresponding contribution to the intellectual ability dimension. In addition, JH showed modest validity coefficients for the counselor and work supervisor criterion ratings as well as reasonably acceptable estimates of reliability. Some possible difficulty is worth noting in the item characteristics for the female sample, wherein item mean scores tend to be overly large (i.e., females tend to give the most "desirable" or socially acceptable response choices on almost all of the 11 items). Thus, extension of the scaled response choices to four or five alternatives, as opposed to the present three, should be considered.

Motivation for Vocational Achievement: The separate analyses of the 11 items of MVA:I and the seven items (with their differing format) that make up MVA:II clearly indicate that MVA:I, as an item-format approach, results in a far superior measure. On the basis of the pattern of interrelationships within the battery for MVA:II, and its correlations with the criterion ratings, the measure appears to have little promise in any future development or application efforts. MVA:I, on the other hand, is a measure with a strong influence throughout the battery especially for females and, in general, has the best validity coefficients of any of the attitudinal measures. Notable are the striking correlations of MVA:I with the Job Holding Skills measure ($r = .49$ for the males and .44 for the female sample) as well as a substantial degree of positive relationship (among females) with the Clerical and Service occupational clusters of the "Interest in Vocational Tasks" measure (r = .48 and .32 respectively). Item characteristics and reliability estimates further reinforce the suitability of this measure as an instrument for use with N.Y.C. Out-of-School enrollees.

Interest in Vocational Tasks: The seven occupational clusters of four items each that make up this measure formed, essentially, their own unique factor of vocational interest and showed relatively few significant intercorrelations with other measures of the battery. The pattern of factor loadings and those few significant intercorrelations with other measures that were found are of value in deciding about the further development of any particular vocational cluster. For example, the somewhat unique aspects of the Outdoor cluster--that have already been considered--might lead to dropping of this group of items from the VI measure since it seems too heavily influenced by the perceived status of the jobs (i.e., low level, menial tasks) rather than by task interest *per se*. Clerical, Business and Service task clusters were found to be of particular importance for the female interest patterns; while Technical, Science and Business clusters among males provided almost all of the significant intercorrelations with the other measures of the battery. Validity for any of the VI clusters is almost nil with one exception--the significant correlation of the Service cluster for females with counselor ratings (r = .25). The particularly small number of items that comprise each occupational cluster, along with the imperfections in the criterion scales, would not lead one to expect a pattern of extensive intercorrelations between these two sets of measures.

With regard to future application of the VI pictorial-verbal measure a question arises of the suitability of the occupational clusters used here for scoring purposes. The problem is one of whether the grouping of items based upon the all-verbal California Occupational Preference Survey (Knapp, Grant, & Demos, 1966) (as derived from a primarily middle-class sample of students and factor analysis of the data) is optimum for a group of disadvantaged adolescents. The question is one of whether the perception of occupational tasks in a disadvantaged adolescent, out-of-school group, differs in ways that might suggest different clusters for scoring purposes. This is especially relevant where pictorial information is added to the verbal content of a measure and may lend different meanings to any occupational tasks presented. Differences may also be engendered by the extent to which a respondent depends upon a verbal or pictorial information mode when both are provided. Accordingly, the 28-item male and female VI scales were factor analyzed, using a principal components solution. Eight factors were extracted from each 28 x 28 item matrix and subjected to a varimax rotation. For each solution five factors are interpretable. Generally, only four or five loadings reached magnitudes of a reasonable level on each factor and the loadings of these items on the factors are shown in Table 9 for males and Table 10 for females.

Factor I for males is best termed a High Level Technical-Business cluster with a strong focus on the level of the skills required rather than the particular application of the task (i.e., a highly skilled scientific task tends to be seen in the same way as a highly skilled business or technical task).

Table 9

Factors and Factor Loadings for Items of the VI Measure (Males, N=123)

Item	Loading
Factor I--High Level Technical and Business	
Give advice on how to design a new airplane.	.67
Write up reports for a company on how much money it's making.	.69
Read dials and adjust switches in a research laboratory.	.74
Take charge of the building of a large bridge.	.57
Factor II--Low-Level (Outdoor)	
Feed cows, take care of them and milk them.	.83
Plant lawns, shrubs and bushes and take care of them.	.73
Do truck garden farming.	.78
Do things for sick people in hospitals to make them comfortable.	.44
Factor III--Office-Clerical	
Run an office machine that handles all kinds of paperwork.	.67
Sterilize instruments used in a laboratory.	.52
Be responsible for taking care of office files.	.69
Sort mail in a post office.	.48
Give advice on how to design a new airplane.	.42
Factor IV--Technical Skills	
Inspect machinery.	.70
Cut and join metal with a welding torch.	.78
Fix wires on telephone poles.	.49
Put plaster on walls and ceilings.	.46
Take charge of the building of a large bridge.	.42
Factor V--Personal Service	
Help poor families that have problems.	.78
Be a department manager in a store.	.50
Sell automobiles to people.	.46
Help sick people in hospitals keep busy by showing them how to make things.	.43
Help people make plans to go on trips.	.38

Table 10

Factors and Factor Loadings for Items of the VI Measure (Females, N=133)

Item	Loadings
Factor I--Clerical	
Be responsible for taking care of office files.	.81
Run an office duplicator or copy machine.	.75
Keep records of stock and supplies.	.78
Write up financial reports for a company.	.68
Handle money that people put in and take out of the bank.	.61
Factor II--Low-Level (Outdoor)	
Plant and care for lawns, shrubs and flowers.	.84
Raise and take care of bees.	.75
Feed cows, take care of them and milk them.	.58
Cut and join metal with a welding torch.	.53
Run a pressing machine.	.42
Factor III--Low-Level Technical Skills	
Bake bread in large oven.	.74
Sell cosmetics to women at their homes.	.70
Run a pressing machine.	.62
Set up artificial flowers in a store display.	.54
Fix jewelry and watches.	.45
Factor IV--High-Level Technical	
Do illustrations for a magazine.	.75
Collect information about weather conditions.	.63
Sell houses to people.	.50
Cut and join metal with a welding torch.	.50
Help people make travel arrangements.	.50
Factor V--Personal-Service	
Help sick people in hospitals keep busy by showing them how to make things.	.85
Do things for sick people in hospitals--like adjusting beds to make them comfortable.	.77
Do hair styling.	.58
Help poor families that have problems.	.54
Factor VI--Laboratory-Technician	
Sterilize laboratory instruments.	.79
Wash and set up laboratory glassware used in scientific experiments.	.76
Read dials and adjust switches in a research laboratory.	.38

Factor II is unmistakably similar to the Outdoor cluster, but also picks up an additional lower level occupation (item #13), in support of the contention that the perception of these occupations is not on the basis of their "outdoor" quality but their menial level. Thus, it is designated as Low-Level (Outdoor) Tasks.

Factor III is identified fairly readily as an Office-Clerical factor. Of interest is that the sterilizing of instruments is seen in the Office-Clerical context rather than in a laboratory-technical setting. The possibility is suggested that the picture and task statement are confusing or ambiguous to the male enrollee. The most easily defined factor of the group is Factor IV which is termed a Technical Skills factor and along with Factor III constitutes the two clusters closest to the ones derived for the COPS measure.

The major item defining Factor V (i.e., the highest loading) is the one that pictures an individual most appropriately considered as a social worker, talking to a family (item #21). The remaining items fit a pattern of job tasks that involve dealing with people directly and helping them or selling them something. The dimension seems best considered as a Personal Service factor.

For females, the group of factors is markedly similar to those defined for males. Thus, in Table 10 Factor I is seen to be a Clerical cluster, Factor II a Low-Level Occupational (Outdoor) cluster and Factor V a Personal-Service cluster (in direct correspondence to the male sample). The Low-Level Technical Skills (Factor III) and the High-Level Technical Skills clusters are also counterparts in some respects to the High-Level Technical and Business and Technical-Skills clusters found for males. The additional factor (Factor VI) for females is unique and is designated as a Laboratory Technician cluster, since it obviously loads on several tasks that are carried out in a laboratory setting.

Further development of items and grouping of tasks into occupational clusters for scoring purposes should take into account the groupings suggested by these factors rather than relying exclusively upon clusters derived from the California Occupational Preference Survey, or any measure developed with a middle-class sample.

Vocational Aspirations and Plans: From the item characteristics of these two measures (regardless of the underlying status continuum) it is evident that item means and variances show reasonable patterns of response. However, use of the measures for analysis purposes was dependent upon summary scores of some sort. In this case, the built-in status continuum with item pairs at eight status levels and the use of identical occupations for the VA and VP items, allowed for the derivation of the three scores utilized--i.e., discrepancy of occupational choice over the status pairs (VA and VP Discrepancy scores) and the difference between degree of occupational desirability in terms of aspirations and plans (VA-VP Difference score). In their relationship to other measures of the battery and the criterion ratings, their role was not an outstanding one. Primarily, there was no significant validity achieved for any of the three scores (although the correlations of these scores with the criteria almost all show a tendency toward a negative relationship). Among the males the VA-VP Difference score, which contributed significantly to interpretation of several of the factors, is related to a number of measures in the battery in a logical way and would appear to hold some promise for use with this group. For females, on the other hand, the VA Discrepancy score seemed to play a more significant role in terms of its relations to other measures.

Reliability of the two measures (split-half comparisons based upon matched item-pairs at each status level) was found to be acceptable for a scale with so few items and to indicate a fair degree of consistency in the perception, by enrollees, of the underlying status continuum.

Further evidence of logical responses to the two measures can be obtained from more detailed analyses. Examination of the present data can be based upon logical hypotheses and on the basis of prior research findings. Thus responses on the Vocational Aspirations and Vocational Plans measures would be expected to show generally higher scores for aspirations than for plans (Empey, 1956). Item means should also show a progressive increase with an increase in underlying status level--especially for aspirations.

The item means for Vocational Aspirations and Plans are presented in Table 11.

Table 11

Vocational Aspirations and Plans: Mean Scores by Status Levels

	Male		Female	
	Aspirations	Plans	Aspirations	Plans
Status Level	Mean	Mean	Mean	Mean
1	1.47	1.64	1.49	1.33
2	2.18	2.08	2.09	1.74
3	2.28	2.09	2.37	1.75
4	2.62	2.27	2.57	2.27
5	2.94	2.66	2.92	2.54
6	2.98	2.40	3.31	2.65
7	2.66	1.98	3.26	2.31
8	3.22	2.01	2.74	1.65

Aspirations are found, as one might expect, to be consistently higher than vocational plans for both males and females. But the point of interest is the progressive increase in mean values for aspirations with an increase in status level. Obviously occupational desirability, in terms of aspirations, is tied closely to a status continuum for this group. For vocational plans, however, the males show a distinct downward shift (toward greater "reality"?) with the highest mean value at status level 5 (Plumber; Auto Mechanic) while there is less of a change for females in this respect. For the female sample the highest mean values for plans and aspirations tend to occur at about the same status level--i.e., level 6 which is defined by the occupations of "Secretary" and "Nurse." The results might be interpreted to mean that females simply express job aspirations that are closer to their plans. However, there is a possible artifact in the status scale which prevents a more orderly progression of female aspirations from low to high status choices. The suggested difficulty lies in the general lack of occupations at the highest status levels that tend to be perceived as customarily occupied by women (e.g., judge, lawyer, engineer) and is equally true for the two occupations finally selected for the present measures--"Doctor" and "Scientist." Thus, the unusually sharp reversal in status perception of both plans and aspirations for the mean values of status-level 8 might result from such occupations being viewed as totally outside the realm of possibility by females.

V. CONCLUSIONS AND RECOMMENDATIONS

Any conclusions or broad generalizations drawn from results based upon the initial use of a battery of some 13 different evaluation measures (and two criterion scales) should be considered with qualification and caution. Exceptions to generalizations will not be difficult to find for an individual measure, or in some limited phase of the data analysis. Nevertheless, keeping in mind the purposes of this study and, particularly, its preliminary or exploratory nature, there are a number of reasonable conclusions that follow from the results.

(1) First, and as an overall conclusion, it would appear that the battery of measures shows potential value for evaluation purposes (and possibly other purposes) with a disadvantaged, adolescent group that possesses relatively low verbal-reading levels. In general, the measures would seem to hold somewhat more promise for use with the N.Y.C. out-of-school male enrollees than with the females. Although, on the basis of the present data, a number of the measures can be considered equally useful for both groups. In essence, when the content of the measures is consonant with the limitations in verbal skills and culturally relevant to the group being evaluated, the pattern of responses and the resulting statistical indices are logical and interpretable.

(2) As to more specific conclusions: despite the homogeneity of the enrollee sample (primarily in terms of cultural background and race) the characteristics of the items used for these measures were found to be acceptably adequate in terms of mean scores, variability and item-total-test correlations. When the test material is meaningful, adequate variation between individuals is possible from a so-called "homogeneous" sample.

(3) Similarly, the reliability (internal consistency) estimates for the various measures were at levels that indicate sufficient consistency of response to allow for multi-purpose utilization of the battery (i.e.,

for prediction of future performance to some extent and for guidance and placement purposes, as well as for evaluation). Such consistency lends further support to any contention that the items of these measures do communicate their content or meaning to a reasonable degree. Findings of adequate reliability are of particular importance here, since they suggest the possibility of maintaining the relatively short test length desirable for a disadvantaged group.

(4) An attempt to determine validity for each measure, based upon criteria consisting of counselor and work supervisor ratings, results in rather low to moderate (but statistically significant) relationships between a number of the measures and these evaluations. In addition, the pattern of these relationships is fairly logical. However, the criterion evaluations themselves are open to question in certain respects since they are highly generalized (i.e., they indicate a large "halo" effect) and produce overly lenient judgments, especially for the female enrollees. Changes in instructions and administration of the rating scales are felt to be required if they are considered for future use as criteria. Furthermore, counselors and work supervisors do not tend to agree particularly well in their evaluations of what constitutes a "suitable" or "unsuitable" enrollee. Obviously, there is the unanswered question of the utility of counselor and work supervisor ratings as criteria and whether these evaluations, in turn, bear some reasonable relation to later enrollee vocational "success." If it is assumed, for purposes of the present study, that these ratings constitute acceptable criteria, then the pattern of validity coefficients found would indicate that a number of measures in the battery show moderate predictive validity.

(5) The pattern of intercorrelations between the measures of the battery are readily interpretable (implying a degree of construct validity) and reinforce the general findings of response consistency on the part of the enrollees. The factors, or groupings of intercorrelated measures, that help to define the underlying dimensions being measured, are found to be confined largely to an intellectual skills area of performance (i.e., measures requiring particular knowledge or some degree of verbal reading skill) and several attitudinal dimensions. However, there is some degree of overlap for these, such that certain of the attitudinal measures and intellectual-ability measures are somewhat related.

When the factors that make up the battery are considered along with the criterion ratings, it is clear that the male enrollees are evaluated (by counselors and work supervisors) primarily on the basis of intellectual capabilities (i.e., measures that require "reasoning" or following verbal directions) whereas females tend to be evaluated on the basis of an attitudinal conformity component (primarily in the work situation). Thus, certain of the measures might be expected to perform differentially as predictors of later success for the male and female enrollees.

(6) Of the measures in the battery, the most effective single one--for those categorized as intellectual-ability measures was the test designated as "Job Seeking Skills"--which required interpretation of employment want-ads and portions of a job application blank. Of the attitudinal measures the most promising possibilities are shared by the measures designated as "Motivation for Vocational Achievement" and "Job Holding Skills" which deal with feelings about the importance of working and how one should behave on a job.

Future development or utilization of the instruments in this battery undoubtedly require additional understanding of their characteristics and certain cautions in their application. Recommended areas of future concern, or needed research effort, involve:

(1) First, the development of normative data on a larger and more diverse sample of disadvantaged adolescents in Neighborhood Youth Corps projects. This seems essential, if for no other purpose than to determine an individual's "relative deprivation in comparison with others who have also been denied good homes, good neighborhoods, good diets, good schools and good teachers" (SPSSI, 1964). Thus, it should be possible to compare performance for an enrollee on any given measure in relation to scores made by some pertinent reference groups of disadvantaged adolescents (based upon sex, age, geographic location, or other desired subgroupings). As part of data collection for normative uses, the wider application of the measures and the opportunity for reanalyses of resulting data serves, of course, to add to their status and utility as formal assessment instruments.

Clearly, appropriate normative data should be obtained from individuals with a reasonable comprehension of English, since it is not feasible to use the present measures with those who have only minimal capability with the spoken as well as written language. (Translation to Spanish for use with Puerto Rican or other Spanish-speaking enrollees is not precluded for later development.) It is also questionable as to whether the present battery would be applicable to such unique cultural subgroups as American Indians.

(2) As a second area of future concern in the use of the measures, there remains a particular problem involving group administration for those

enrollees at an exceptionally low-verbal reading level (e.g., fourth grade or lower). Separate administration in small groups may be required for those enrollees who seem to profit from the oral presentation; whereas, for those at relatively high verbal reading levels (e.g., seventh grade level or higher) the examiner's oral presentation can be annoying and they tend to move ahead of the group in their responses to the items. Some screening, based upon a set of trial items or background information (where available) should be considered for this purpose. Conceivably, tape recorded instructions--controllable by the enrollee--could be utilized to deal with the problem.

(3) Development of alternate forms of the measures constitutes a third area for future consideration. This is of importance if the measures are to serve as evaluative instruments for detecting change in enrollee characteristics over a relatively short period of enrollment in N.Y.C. Thus, the alternate forms should serve to overcome any effects of recall of the material in an evaluation study and also provide a means for obtaining test-retest reliability estimates over relatively short time periods (e.g., several days).

(4) Another major point to be considered in future application of the evaluation measures is the development of relevant, post-N.Y.C. criterion measures of enrollee vocational adjustment. Suitable criteria are necessary to provide evidence regarding the characteristics that are being assessed by any battery of measures. Validation of the measures against longer-range criteria than ratings by counselors or work supervisors would also provide needed evidence of their accuracy and suitability for a variety of potential uses by N.Y.C.

Wider application of the measures in the context of a youth-work training program are recommended, generally, as the only means to determine which measures are most effective and for which situations. Therefore, the process of developing the measures should be a continuing one, during which their utilization as evaluative or guidance instruments adds to the knowledge of their applicability and also provides some uniformity, or consistency, of research results. Continued use of "ad lib," informal questionnaires, as so-called "measures," with psychometric properties that are largely unknown and content that meets the changing needs as perceived by each individual investigator, can only provide self-defeating inconsistencies in conducting a program of evaluation and research.

REFERENCES

Ammons, R. B., Butler, M. N., & Herzig, S. A. The Vocational Apperception Test: Advanced form. Missoula, Montana: Psychological Test Specialists, 1949.

Anastasi, A. Culture fair testing. Education Digest, 1965, 30, 9-11.

Barnette, W. L., Jr., & McCall, J. N. Validation of the Minnesota Vocational Interest Inventory for vocational high school boys. Journal of Applied Psychology, 1964, 48, 378-382.

Bernberg, R. E. Human Relations Inventory. Chicago: Psychometric Affiliates, 1954.

Campbell, J. Testing of culturally different groups. College Entrance Examination Board, Research and Development Report 63-4, No. 14. Princeton, N. J.: Educational Testing Service, 1964.

Clark, C. A. The use of separate answer sheets in testing slow-learning pupils. Journal of Educational Measurement, 1968, 5 (1), 61-64.

Clark, K. E. Vocational interests of nonprofessional men. Minnesota: University of Minnesota Press, 1961.

Clark, K. E., & Campbell, D. P. Minnesota Vocational Interest Inventory. New York: Psychological Corporation, 1965.

Coleman, J. S. Equality of educational opportunity. U. S. Dept. of Health, Education, and Welfare, Office of Education. Washington, D. C.: U. S. Government Printing Office, 1966.

Davis, A., & Dollard, J. Children of bondage. Washington, D. C.: American Council on Education, 1940.

Demos, G. D. Manual for the Demos D-Scale. Beverly Hills, Cal.: Western Psychological Services, 1965.

Deutsch, M., & Brown, B. Some data on social influences in Negro-white intelligence differences. Journal of Social Issues, 1964, 20 (2), 24-35.

Dreger, R., & Miller, K. Comparative psychological studies of Negroes and whites in the United States. Psychological Bulletin, 1960, 57, 361-402.

Dubner, N. A Poverty Youth Program: Towards a need for creating a meaningful testing program. New York: Mobilization for Youth, 1965.

Duncan, O., Hatt, P., & North, C. Occupations and social status. New York: The Free Press of Glencoe, 1961.

Ebel, R. L. The value of internal consistency in classroom examinations. Journal of Educational Measurement, 1968, 5 (1), 71-73.

Educational Design, Inc., & National Association of Manufacturers. How to get a job. Reading, Mass.: Addison-Wesley, 1966.

Educational Testing Service. Specifications for Neighborhood Youth Corps evaluation measures. Princeton, N. J.: 1967. (Multilith report on file at ETS library.)

Elias, G. Elias family opinion survey. Chicago: Psychometric Affiliates, 1952.

Empey, L. T. Social class and occupational aspiration: A comparison of absolute and relative measurement. American Sociological Review, 1956, 21, 703-709.

Feingold, S. N., & List, H. How to get that part time job. New York: Arco Publishing Co., 1958.

Geist, H. The Geist Picture Interest Inventory: Male and female. Beverly Hills, Cal.: Western Psychological Services, 1964.

Ginzberg, E., Ginsburg, S. W., Axelrad, S., & Herma, J. C. Occupational choice. New York: Columbia University Press, 1951.

Gordon, E. W. A question of culture. American Child, 1963, 45, 11-14.

Gough, H. G. California Personality Inventory. Palo Alto, Cal.: Consulting Psychologists Press, 1957.

Gough, H. G. Theory and measurement of socialization. Journal of Consulting Psychology, 1960, 24, 23-30.

Graham, W. R., & Orr, D. B. Development of a listening comprehension test to identify college potential among the disadvantaged. Final Report. Silver Spring, Md.: American Institutes for Research, 1966.

Greene, E. B. Measurements of human behavior. (Rev. ed.) New York: Odyssey Press, 1952.

Guion, R. M. Employment tests and discriminatory living. Industrial Relations, 1966, 5, 20-37.

Guion, R. M. Personnel selection. Annual Review of Psychology, 1967, 18, 191-216.

Herman, M., & Sadofsky, S. Youth-work programs. New York: Center for the Study of Unemployed Youth, Graduate School of Social Work, New York University, 1966.

Holland, J. L. Preliminary manual for Holland Vocational Preference Inventory. Palo Alto, Cal.: Consulting Psychologists Press, 1959.

Horn, J. L. Is it reasonable for assessments to have different psychometric properties than predictors? Journal of Educational Measurement, 1968, 5 (1), 75-77.

Kahl, J. A. Some measures of achievement orientation. American Journal of Sociology, 1965, 70 (6), 669-681.

Kaiser, H. F. The varimax criterion for analytic rotation in factor analysis. Psychometrika, 1958, 23, 187-200.

Karlsen, B., Madden, R., & Gardner, E. Adult Basic Learning Examination. New York: Harcourt, Brace & World, 1967.

Karp, J. M., & Sigel, I. Psychoeducational appraisal of disadvantaged children. Review of Educational Research, 1965, 35 (5), 401-412.

Kirkpatrick, J., Ewen, R., Barnett, R., & Katzell, R. Differential selection among applicants from different socioeconomic or ethnic backgrounds. Final report. New York: Department of Psychology, New York University, May 1967.

Knapp, R. R., Grant, B., & Demos, G. D. California Occupational Preference Survey. San Diego: Educational and Industrial Testing Service, 1966.

Kuder, G. F. Kuder Preference Record. Chicago: Science Research Associates, 1956.

Lennon, R. T. Testing and the culturally disadvantaged child. New York: Harcourt, Brace & World, 1964.

Lopez, F. M., Jr. Current problems in test performance of job applicants: 1. Personnel Psychology, 1966, 19, 10-18.

Magoun, F. A. Successfully finding yourself and your job. New York: Harper & Brothers, 1959.

Mischel, W. Preference for delayed reinforcement and social responsibility. Journal of Abnormal and Social Psychology, 1961, 62, 1-7.

Mobilization for Youth. Intake interview. New York: Unpublished manuscript, 1964.

New York City Commission on Human Rights. *Testing human potential, new techniques for selecting employees from minority groups*. Conference held in New York City, April 25-26, 1968.

Parnicky, J., Kahn, H., & Burdett, A. Vocational interest and sophistication assessment. In J. Parnicky & H. Kahn (Eds.), *Evaluating and developing vocational potential of institutionalized retarded adolescents*. Division of Mental Retardation, New Jersey Department of Institutions and Agencies, 1963.

Rogers, C. R. *Personal adjustment inventory*. Beverly Hills, Cal.: Western Psychological Services, 1961.

Rosenberg, M. *Society and the adolescent self-image*. Princeton, N. J.: Princeton University Press, 1965.

Rosenzweig, S. *Picture frustration study*. Copyright 1948 by S. R. Rosenzweig.

Schneider, L., & Lysgaard, S. The deferred gratification pattern: A preliminary study. *American Sociological Review*, 1953, *18*, 142-149.

Shimberg, B. Review of 1964 revision of Geist Picture Interest Inventory. In O. K. Buros (Ed.), *Sixth Mental Measurements Yearbook*. Highland Park: Gryphon Press, 1965. P. 1954.

Slotkin, H., & Forlano, G. *New programs for dropouts in New York City*. New York: Board of Education, 1962.

Society for the Psychological Study of Social Issues (SPSSI). *Guidelines for testing minority group children*, 1964, *20* (2), 127-145.

Stephenson, R. Occupational aspirations and plans of 443 ninth graders. *Journal of Educational Research*, 1955, *49*, 27-35.

Straus, M. Deferred gratification, social class and the achievement syndrome. American Sociological Review, 1962, 27 (3), 326-335.

Strong, E. K. Strong Vocational Interest Blank for Men. Stanford: Stanford University Press, 1945.

Thorndike, E. L. Chapter 15 in E. F. Lindgurst (Ed), Educational measurement. Washington, D. C.: American Council on Education, 1951.

Thorndike, E. L., & Lorge I. The teacher's word book of 30,000 words. New York: Bureau of Publications, Teacher's College, Columbia University, 1944.

Tucker, J. F. The first 50,000 Neighborhood Youth Corps enrollees. Washington, D. C.: U. S. Department of Labor, Bureau of Labor Statistics, 1966.

U. S. Department of Labor, Public Employment Service. How to get and hold the right job. Washington, D. C.: U. S. Government Printing Office, 1964.

U. S. Department of Labor, Manpower Administration. Handbook for sponsors, 1967. Washington, D. C.: U. S. Government Printing Office, 1967.

U. S. Employment Service. Dictionary of occupational titles. Washington, D. C.: U. S. Government Printing Office, 1965.

U. S. Office of Economic Opportunity. Occupational training manual. (Rev. ed.) Washington, D. C.: U. S. Government Printing Office, 1966.

U. S. Office of Economic Opportunity. Economic Opportunity Act of 1964, as amended. Washington, D. C.: U. S. Government Printing Office, 1967.

U. S. Senate Committee on Labor and Public Welfare. Profile of Youth, 1963. Washington, D. C.: U. S. Government Printing Office, 1964.

Varo, R. D. Career aspirations of Negro and Puerto Rican youth. Negro Educational Review, 1964, 15, 60-62.

Walther, R., & Magnusson, M. L. A retrospective study of the effectiveness of out-of-school Neighborhood Youth Corps programs in four urban sites. Washington, D. C.: Social Research Group, The George Washington University, 1967.

Weingarten, K. P. Picture Interest Inventory. Los Angeles, Cal.: Test Bureau, 1958.

Wesman, A. G. Intelligent testing. American Psychologist, 1968, 4, 267-274.

Wylie, R. The self-concept: A critical survey of pertinent research literature. Lincoln: University of Nebraska Press, 1961.

Appendix A

Item Data for Enrollee Measures

(Means, Variances and Item-Total Test Correlations)

Job Knowledge

	Male						Female					
	Rural (N≅29)			Urban (N≅93)			Rural (N≅30)			Urban (N≅103)		
Item	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r
1)	.34	.48	.60	.45	.50	.34	.87	.35	.63	.97	.17	.19
2)	.48	.51	.24	.55	.50	.13	.87	.35	.34	.93	.25	.35
3)	.73	.45	.56	.73	.45	.44	.57	.50	.57	.56	.50	.37
4)	.93	.26	.25	.91	.28	.44	.83	.38	-.20	.87	.33	.03[a]
5)	.96	.19	.29	.98	.15	.33	.57	.50	.20	.47	.50	.35
6)	.89	.32	.19	.91	.29	.51	.86	.35	.60	.84	.37	.53
7)	.69	.47	.48	.87	.34	.49	.87	.35	.65	.96	.19	.26
8)	.97	.19	.42	.91	.29	.46	.90	.31	.33	.96	.19	.18
9)	.66	.48	.12	.70	.46	.31	.70	.47	.62	.80	.40	.58
10)	.69	.47	.18	.47	.50	.15	1.00	-.00	-.00	.98	.14	.18[a]
11)	1.00	-.00	-.00	.99	.10	.15[a]	.83	.38	.60	.86	.35	.41
12)	.93	.26	.27	.92	.27	.41	.83	.38	.51	.88	.32	.53
13)	.62	.49	.19	.60	.49	.36	.97	.18	.50	.97	.17	.33
14)	.83	.38	.79	.88	.33	.64	.97	.18	.36	1.00	-.00	-.00[a]
15)	.79	.42	.15	.82	.39	.30	.70	.47	.58	.74	.44	.47
16)	1.00	-.00	-.00	.97	.18	.37[a]	.93	.25	.01	.86	.34	.26
17)	.66	.48	.11	.57	.50	.25	.73	.45	.37	.71	.46	.51
18)	.48	.51	.03	.40	.49	.18	.67	.48	.59	.88	.32	.42
19)	.59	.50	.70	.71	.46	.46	.90	.31	.52	.80	.40	.31
20)	.72	.45	.09	.71	.46	.37	.63	.49	.76	.69	.46	.61
21)	.90	.31	.66	.93	.25	.54	.67	.48	.60	.67	.47	.43
22)	1.00	-.00	-.00	.98	.15	.21[a]	1.00	-.00	-.00	.99	.03	.25[a]
23)	.90	.31	.17	.92	.27	.21	.53	.51	.36	.65	.48	.27
24)	.83	.38	.65	.83	.38	.43	.63	.49	.20	.69	.47	.18
25)	.97	.19	.42	.94	.25	.52	.93	.25	.07	1.00	-.00	-.00[a]
26)	.83	.38	.40	.92	.27	.21	.33	.48	.31	.58	.50	.48
27)	.90	.31	.17	.94	.25	.06[a]	.90	.31	.60	.99	.10	.16
28)	.55	.51	.28	.71	.46	.50	--	--	--	--	--	--
29)	.45	.51	.29	.45	.50	.37	--	--	--	--	--	--
30)	.66	.48	.41	.56	.50	.24	--	--	--	--	--	--

---

[a] Items eliminated from analyses.

Vocational Aspiration

	Male						Female					
	Rural (N=29)			Urban (N=93)			Rural (N=30)			Urban (N=103)		
Item	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r
1)	2.55	.69	.26	2.58	.81	.47	3.00	.69	.39	2.90	.76	.45
2)	2.59	.95	.45	2.55	.95	.32	3.20	.85	.09	2.88	.88	.26
3)	3.14	.74	.52	2.58	.92	.39	2.63	.85	.62	2.51	.87	.45
4)	2.41	.78	.25	2.55	.94	.33	3.60	.67	.51	3.46	.84	.28
5)	3.24	.79	.77	3.12	.90	.44	2.53	1.01	.67	2.56	1.04	.45
6)	3.31	.89	.59	3.29	.87	.40	3.07	.69	.09	2.85	.94	.44
7)	1.31	.60	.47	1.53	.75	.48	1.60	.97	.40	1.26	.52	.38
8)	2.93	.70	.39	2.59	.73	.30	3.43	.68	.18	3.31	.66	.43
9)	2.83	1.07	.66	2.74	.99	.42	1.40	.56	.65	1.55	.76	.38
10)	3.10	.77	.69	2.27	.83	.52	1.83	.79	.72	1.58	.68	.41
11)	2.76	.79	.31	2.51	.96	.58	2.30	1.02	.73	2.17	.85	.55
12)	2.38	.68	.32	2.00	.88	.39	2.37	.56	.08	2.43	.82	.52
13)	2.24	.83	.27	1.73	.82	.34	2.97	.61	.57	2.65	.68	.50
14)	3.48	.63	.39	3.24	.93	.45	2.83	.70	.42	2.62	.85	.62
15)	1.76	.58	.07	1.38	.59	.46	3.33	.84	.48	3.06	.87	.55
16)	3.17	.85	.65	3.26	.87	.36	3.53	.68	.14	3.09	.85	.52

Vocational Plans

	Male						Female					
	Rural (N≅29)			Urban (N≅93)			Rural (N≅30)			Urban (N≅103)		
Item	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r
1)	2.14	.69	.43	1.94	1.02	.51	1.80	1.00	.61	1.72	.78	.48
2)	1.90	.72	.39	1.96	.98	.45	2.63	.96	.32	2.55	.97	.30
3)	2.72	.80	.44	2.15	.98	.52	2.20	.81	.43	1.79	1.00	.43
4)	2.38	.78	-.00	2.34	1.08	.60	3.07	1.01	.38	2.99	1.11	.15
5)	1.90	1.05	.67	1.96	1.05	.42	1.60	.86	.47	1.57	.81	.46
6)	1.83	.93	.52	2.15	1.09	.35	2.90	.92	.35	2.41	1.03	.40
7)	1.66	.77	.08	1.70	.95	.38	1.67	1.12	.37	1.19	.54	.29
8)	2.62	.82	.25	2.54	1.02	.54	2.70	.92	.53	2.45	.95	.51
9)	2.10	1.05	.63	2.00	1.04	.54	1.27	.52	.27	1.22	.56	.50
10)	2.66	.86	.62	2.19	1.05	.60	1.60	.86	.10	1.28	.63	.48
11)	2.66	.81	.32	2.27	1.05	.42	1.87	1.04	.42	1.55	.75	.39
12)	2.10	.72	.49	1.80	1.00	.55	2.07	.78	.14	2.05	.89	.47
13)	2.14	.83	.16	1.72	.96	.53	2.70	.79	.71	2.43	.94	.40
14)	2.83	1.04	.45	2.99	1.05	.42	2.63	.93	.20	2.14	.98	.47
15)	2.14	.92	.36	1.43	.76	.42	2.53	1.14	.52	2.22	1.07	.47
16)	2.43	1.07	.37	2.58	1.05	.50	2.50	1.17	.52	1.99	.98	.55

Vocational Interest

	Male						Female					
	Rural (N=29)			Urban (N=94)			Rural (N=30)			Urban (N=103)		
Item	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r
1)	2.90	.82	.71	2.87	.80	.26	2.80	.76	.46	2.82	.83	.38
2)	2.76	.87	.42	2.23	.92	.51	2.57	.97	.26	2.72	1.04	.48
3)	3.00	.76	.45	2.95	1.01	.46	2.57	1.01	.50	2.47	.94	.51
4)	2.62	1.05	.57	2.57	.96	.40	2.37	.96	.48	2.16	.89	.48
5)	2.03	.91	.41	1.57	.93	.10	1.90	.96	.43	1.68	.81	.54
6)	2.90	.77	.55	2.52	.92	.53	1.87	.82	.52	1.98	.86	.64
7)	3.31	.76	.48	3.11	.80	.53	2.20	1.10	.44	2.01	.90	.57
8)	2.72	.92	.27	2.96	1.00	.59	3.20	1.00	.35	3.15	.87	.44
9)	3.10	.77	.51	3.12	.91	.52	2.80	1.00	.56	2.87	.81	.54
10)	2.55	.87	.33	2.09	.95	.54	2.87	.82	.04	2.81	.86	.47
11)	2.00	.85	.05	1.52	.83	.11	1.57	.77	.57	1.40	.69	.50
12)	3.07	.70	.47	2.76	.85	.42	3.47	.63	.27	3.19	.75	.40
13)	3.00	1.00	.64	2.25	.97	.45	2.87	.94	.44	2.82	.86	.44
14)	3.00	.60	.29	2.81	.82	.51	2.93	.74	.11	2.65	.81	.40
15)	2.93	1.03	.66	2.75	.91	.43	1.63	.72	.64	1.77	.94	.67
16)	3.00	.71	.58	2.61	.85	.50	2.20	.92	.15	1.84	.93	.53
17)	2.86	.92	.46	2.57	.99	.58	1.40	.62	.54	1.38	.73	.31
18)	2.21	.86	.13	1.83	.93	.23	1.13	.35	.26	1.14	.51	.41
19)	2.97	1.02	.77	2.40	1.04	.43	1.27	.58	.50	1.42	.77	.51
20)	2.86	.88	.53	2.96	.84	.48	2.20	.89	.24	2.15	.91	.57
21)	3.41	.73	.66	2.95	.95	.45	2.93	.91	.08	2.99	.93	.47
22)	2.52	.78	.32	2.20	1.05	.20	2.73	.98	.14	2.59	1.03	.52
23)	2.90	.86	.45	2.60	1.00	.53	2.63	.89	.62	2.62	.85	.63
24)	2.83	.97	.69	3.06	1.00	.35	1.27	.58	.04	1.16	.52	.26
25)	2.86	.83	.62	2.76	.91	.57	1.83	.75	.68	1.94	.88	.45
26)	2.90	.86	.39	2.81	1.01	.36	2.77	.86	.51	2.51	.89	.55
27)	2.59	1.12	.63	2.71	1.13	.42	2.13	.94	.21	2.19	1.07	.48
28)	3.38	.78	.39	3.02	.82	.38	2.43	.86	.19	2.17	.87	.53

Attitude toward Authority

	Male						Female					
	Rural (N≅29)			Urban (N≅94)			Rural (N≅20)			Urban (N≅103)		
Item	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r
1)	2.97	1.48	.42	3.09	1.62	.24	3.30	1.87	.30	3.85	1.59	.30
2)	2.97	1.74	.51	2.62	1.65	.46	4.05	1.54	.50	4.24	1.33	.41
3)	2.48	1.57	.35	2.94	1.76	.29	3.45	1.82	.58	3.80	1.32	.23
4)	3.41	1.55	.52	2.91	1.65	.49	4.70	.92	.67	4.20	1.56	.55
5)	3.31	1.54	.72	3.49	1.60	.57	4.05	1.47	.59	3.45	1.58	.55
6)	3.69	1.67	.69	4.13	1.41	.58	3.60	1.82	.50	3.43	1.68	.28
7)	2.00	1.48	.53	2.34	1.70	-.07	2.05	1.28	.21	2.35	1.53	.55
8)	2.83	1.81	.51	3.81	1.79	-.01	4.55	1.15	.79	4.11	1.43	.52
9)	4.03	1.24	.76	3.82	1.53	.64	3.90	1.55	.67	3.30	1.67	.40
10)	2.55	1.50	.38	2.24	1.45	.51	2.70	1.89	.28	2.24	1.61	.30
11)	3.90	1.21	.26	3.81	1.48	.62	3.20	1.67	.44	2.84	1.55	.53
12)	3.52	1.43	.31	3.51	1.53	.51	3.30	1.81	.39	3.50	1.62	.65
13)	3.76	1.35	-.06	3.82	1.62	.48						

Self-Esteem

Item	Male						Female					
	Rural (N=29)			Urban (N=94)			Rural (N=30)			Urban (N=103)		
	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r
1)	2.55	.51	.22	2.55	.56	.10	2.63	.49	.53	2.50	.50	.43
2)	2.00	.53	.38	2.01	.52	.13	2.07	.58	.65	2.02	.42	.43
3)	2.03	.50	.27	2.01	.63	.45	2.13	.43	-.04	2.07	.43	.31
4)	2.59	.50	-.12	2.60	.54	.05[a]	2.87	.35	-.06	2.70	.46	.27
5)	2.10	.49	-.03	2.11	.60	.41	1.93	.25	.25	2.01	.36	.09
6)	2.55	.74	.42	2.81	.53	.43	2.97	.18	-.03	2.91	.32	.30
7)	2.52	.63	.20	2.35	.73	.44	2.63	.61	.15	2.62	.51	.44
8)	2.38	.62	.37	2.32	.61	.38	2.70	.54	.57	2.50	.64	.54
9)	2.48	.74	.54	2.47	.73	.33	2.77	.50	.48	2.60	.60	.38
10)	2.24	.51	.15	2.12	.62	.18	2.17	.38	-.28	2.05	.36	.09[a]
11)	2.38	.78	.59	2.46	.77	.49	2.67	.71	.34	2.59	.66	.56
12)	2.62	.62	.05	2.62	.61	.41	2.87	.35	.05	2.60	.51	.33
13)	2.59	.68	.40	2.52	.74	.45	2.73	.45	.20	2.73	.49	.51
14)	2.07	.59	.61	1.91	.56	.46	2.13	.35	.43	1.96	.31	.31
15)	1.93	.65	.49	1.89	.74	.16	2.37	.56	.20	2.18	.65	.38
16)	2.03	.42	.43	2.01	.58	.25	2.10	.31	.23	2.02	.31	.09

---

[a] Item eliminated from analyses.

Deferred Gratification

	Male						Female					
	Rural (N=29)			Urban (N=94)			Rural (N=30)			Urban (N=103)		
Item	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r
1)	4.14	1.06	.37	3.49	1.32	-.04	4.00	1.23	-.09	4.14	1.02	.23
2)	2.66	1.20	.28	3.02	1.36	.60	3.03	1.00	.23	3.09	1.06	.42
3)	2.52	1.21	.46	2.50	1.32	.37	2.47	1.31	.45	2.65	1.23	.30
4)	2.54	1.32	.33	2.64	1.41	.39	2.03	1.22	.31	2.23	1.16	.12
5)	3.54	1.23	.50	3.57	1.16	.63	3.37	1.03	.67	3.58	1.12	.63
6)	3.59	1.18	.51	3.57	1.12	.55	4.00	.69	.33	3.79	1.00	.62
7)	2.93	1.30	.63	3.35	1.18	.64	3.13	1.22	.63	2.84	1.20	.43
8)	3.54	1.08	.56	3.22	1.28	.54	3.67	.92	.40	3.42	1.17	.55
9)	3.59	1.40	.22	3.39	1.19	.14	3.93	1.08	.09	3.50	1.18	.20
10)	3.21	.98	.52	3.22	1.09	.59	3.07	1.07	.53	3.45	1.14	.46
11)	4.00	1.20	.41	3.60	1.36	.17	3.77	1.01	-.00	3.47	1.19	.07[a]
12)	3.28	1.10	.82	3.06	1.14	.35	3.07	1.44	.58	3.31	1.06	.51
13)	3.48	.91	.28	3.13	1.15	.35	3.63	1.22	.32	3.48	1.12	.12
14)	3.79	1.32	.53	3.82	1.15	.18	4.10	1.12	.45	4.11	1.00	.55
15)	4.21	.98	.09	3.68	1.20	.33	4.27	1.08	.27	3.68	1.03	.16
16)	2.62	1.45	.26	2.64	1.25	.29	2.20	1.24	.18	2.67	1.06	.07

---

[a] Item eliminated from analyses.

Job Seeking Skills

	Male						Female					
	Rural (N=29)			Urban (N=93)			Rural (N=30)			Urban (N=103)		
Item	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r
1)	.83	.38	.36	.56	.50	.55	.68	.48	.03	.84	.36	.35
2)	.21	.41	.04	.15	.36	.31	.11	.32	.08	.27	.45	.27
3)	.72	.45	.36	.76	.43	.37	.75	.44	.58	.88	.32	.49
4)	.62	.49	.58	.67	.47	.42	.70	.47	.55	.76	.43	.52
5)	.46	.51	.48	.44	.50	.24	.46	.51	.43	.49	.50	.50
6)	.59	.50	.61	.73	.45	.48	.75	.44	.18	.75	.43	.34
7)	.62	.49	.53	.77	.42	.41	.79	.42	.29	.89	.31	.22
8)	.31	.47	.49	.44	.50	.49	.71	.46	.57	.69	.47	.23
9)	.86	.35	.28	.81	.40	.28	.52	.51	-.19	.78	.42	.27
10)	.14	.35	.55	.39	.49	.62	.37	.50	.46	.41	.49	.51
11)	.79	.41	.51	.75	.43	.49	.61	.50	.56	.62	.49	.34
12)	.90	.31	.43	.88	.33	.25	.96	.19	.34	.96	.20	.16
13)	.90	.31	.50	.91	.28	.46	1.00	.00	-.00	.99	.10	-.05[a]
14)	.79	.42	.50	.81	.40	.57	.86	.35	.58	.95	.22	.39
15)	.21	.41	.35	.21	.41	.45	.40	.50	.54	.27	.45	.42
16)	.45	.51	.39	.58	.50	.57	.67	.48	.66	.79	.41	.52
17)	[illegible]	[illegible]	.60	.57	.44	.41	.51	.47	.35	.49	.50	.48

---

[a] Item eliminated from analyses.

Job Holding Skills

	Male						Female					
	Rural (N=29)			Urban (N=93)			Rural (N=30)			Urban (N=103)		
Item	Mean	Sigma	Item - Total r	Mean	Sigma	Item - Total r	Mean	Sigma	Item - Total r	Mean	Sigma	Item - Total r
1)	2.90	.41	.45	2.82	.47	.46	2.97	.18	.00	2.90	.29	.66
2)	2.52	.57	-.02	2.35	.58	.57	2.33	.55	.56	2.60	.51	.34
3)	2.62	.62	.61	2.61	.53	.18	2.87	.35	.53	2.72	.47	.36
4)	2.82	.54	.69	2.46	.82	.28	3.00	.00	-.00	2.86	.49	.42
5)	2.25	.59	.49	2.22	.68	.48	2.43	.63	.48	2.64	.54	.58
6)	2.82	.48	.52	2.57	.63	.35	3.00	.00	-.00	2.80	.43	.67
7)	2.83	.47	.85	2.76	.58	.49	2.93	.37	.62	2.90	.37	.50
8)	2.93	.37	.33	2.74	.63	.69	2.97	.19	.28	2.91	.35	.37
9)	2.93	.26	.61	2.61	.78	.54	2.97	.18	.37	2.93	.32	.43
10)	2.41	.82	.45	2.38	.82	.46	2.83	.53	.09	2.51	.77	.36
11)	2.69	.47	.49	2.53	.64	.31	2.83	.38	.42	2.80	.47	.56

Motivation for Vocational Achievement

	Male						Female					
	Rural (N≅29)			Urban (N≅92)			Rural (N≅30)			Urban (N≅103)		
Item	Mean	Sigma	Item - Total r	Mean	Sigma	Item - Total r	Mean	Sigma	Item - Total r	Mean	Sigma	Item - Total r
1)	4.38	.82	.26	4.12	1.01	.60	4.17	.87	.60	4.19	1.00	.65
2)	4.17	.89	.46	4.22	.84	.57	4.37	.85	.65	4.28	.90	.40
3)	3.90	1.05	.43	3.54	1.15	.49	3.27	1.23	.66	3.50	1.20	.35
4)	4.21	.94	.33	3.97	1.13	.39	4.47	1.07	.51	4.43	.97	.52
5)	3.83	1.17	.66	4.07	1.10	.46	4.20	.85	.46	4.16	.97	.44
6)	4.55	.87	.50	4.50	.77	.74	4.67	.61	.27	4.51	.77	.59
7)	4.24	.87	.49	4.22	.90	.61	4.30	.99	.47	4.32	.91	.60
8)	3.62	1.21	.75	3.22	1.52	.47	4.21	1.05	.06	3.69	1.29	.36
9)	4.25	.80	.66	4.12	1.05	.71	4.30	.79	-.04	4.13	.97	.46
10)	3.11	1.10	.36	3.26	1.20	.32	3.03	1.30	.31	3.57	1.19	.23
11)	3.45	1.40	.67	3.41	1.30	.71	2.90	1.45	.69	2.64	1.31	.57
12)	4.34	1.04	.66	4.00	1.12	.58	4.23	1.01	.70	4.12	1.00	.57
13)	3.55	1.43	.62	2.86	1.36	.70	2.00	1.31	.57	1.95	1.13	.58
14)	3.24	1.30	.62	2.96	1.43	.53	2.66	1.45	.42	2.85	1.27	.48
15)	4.10	1.14	.64	3.43	1.30	.60	3.10	1.06	.62	3.11	1.09	.58
16)	3.52	1.24	.76	2.88	1.39	.65	2.63	1.25	.71	2.31	1.15	.73
17)	3.79	1.11	.29	3.11	1.46	.61	2.72	1.46	.63	2.88	1.47	.47

Practical Reasoning

Map Reading

	Male						Female					
	Rural (N=29)			Urban (N=90)			Rural (N=30)			Urban (N=103)		
Item	Mean	Sigma	Item - Total r	Mean	Sigma	Item - Total r	Mean	Sigma	Item - Total r	Mean	Sigma	Item - Total r
1)	.86	.36	.38	.75	.43	.53	.86	.35	.22	.68	.47	.48
2)	.86	.35	-.14	.70	.46	.46	.76	.44	.11	.58	.50	.38
3)	.48	.51	.51	.67	.47	.44	.36	.49	.34	.40	.49	.50
4)	.21	.41	.09	.44	.50	.51	.53	.51	.48	.40	.49	.27
5)	.24	.44	.53	.57	.50	.41	.21	.41	.33	.62	.49	.47
6)	.32	.48	.66	.42	.50	.55	.30	.47	.29	.42	.50	.33
7)	.28	.45	.58	.42	.50	.70	.41	.50	.36	.39	.49	.55
8)	.45	.51	.65	.70	.46	.48	.59	.50	.53	.66	.48	.48
9)	.17	.38	.27	.40	.49	.56	.52	.51	.68	.39	.49	.47
10)	.52	.51	.31	.51	.50	.20	.34	.48	.24	.51	.50	.44

Practical Reasoning

Zip Code

	Male						Female					
	Rural (N≅29)			Urban (N≅89)			Rural (N≅29)			Urban (N≅102)		
Item	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r	Mean	Sigma	Item-Total r
1)	.69	.47	.64	.93	.25	.41	.97	.19	.03	.89	.31	.59
2)	.79	.41	.52	.77	.42	.73	.75	.44	.81	.83	.38	.62
3)	.66	.48	.52	.63	.48	.66	.75	.44	.67	.70	.46	.43
4)	.39	.50	.57	.60	.49	.53	.75	.44	.67	.64	.48	.53
5)	.72	.45	.79	.85	.36	.49	.83	.38	.77	.86	.35	.43
6)	.45	.51	.61	.57	.50	.74	.79	.41	.70	.74	.44	.57
7)	.59	.50	.69	.60	.49	.58	.86	.36	.62	.80	.40	.54
8)	.55	.51	.62	.69	.46	.67	.86	.35	.53	.80	.40	.61
9)	.34	.48	.19	.37	.49	.48	.41	.50	.61	.43	.50	.54
10)	.34	.48	.49	.52	.50	.52	.61	.50	.34	.76	.43	.32

Practical Reasoning

File Cards

	Male						Female[a]					
	Rural (N=29)			Urban (N=90)			Rural (N=30)			Urban (N=101)		
Item	Mean	Sigma	Item - Total r	Mean	Sigma	Item - Total r	Mean	Sigma	Item - Total r	Mean	Sigma	Item - Total r
1)	.45	.51	.41	.64	.48	.26	.86	.35	.71	.83	.38	.41
2)	.72	.45	.58	.76	.43	.50	.97	.19	.60	.88	.33	.34
3)	.69	.47	.72	.84	.37	.58	.93	.26	.36	.95	.22	.53
4)	.69	.47	.48	.74	.44	.45	.93	.26	.76	.87	.34	.42
5)	.83	.38	.52	.84	.37	.60	1.00	.00	.00	.94	.24	.60
6)	.62	.49	.68	.80	.40	.47	.86	.35	.32	.93	.26	.54
7)	.59	.50	.38	.78	.42	.50	.89	.32	.38	.93	.26	.70
8)	.66	.48	.72	.76	.43	.55	.86	.35	.54	.90	.30	.56
9)	.59	.50	.38	.61	.49	.73	.75	.44	.59	.84	.37	.66
10)	.66	.48	.51	.71	.46	.64	.79	.41	.57	.93	.26	.60

---

[a] This measure was eliminated completely for females.

Enrollee Rating Scale

	Male						Female					
	Rural (N=29)			Urban (N=91)			Rural (N=30)			Urban (N=103)		
Item	Mean	Sigma	Item - Total r	Mean	Sigma	Item - Total r	Mean	Sigma	Item - Total r	Mean	Sigma	Item - Total r
1)	3.76	1.15	.51	3.81	1.29	.63	4.57	.82	.63	4.17	1.02	.63
2)	3.86	1.33	.58	3.70	1.22	.64	4.57	.73	.74	4.12	1.08	.52
3)	3.28	1.44	.67	2.98	1.53	.55	4.13	1.11	.64	3.57	1.40	.60
4)	3.97	1.45	.64	3.33	1.32	.54	4.30	1.02	.59	3.84	1.10	.49
5)	3.97	1.18	.79	3.90	1.04	.75	4.40	.62	.32	4.48	.62	.46
6)	4.10	1.08	.87	3.88	1.30	.81	4.23	.86	.79	4.14	1.06	.60
7)	4.28	1.00	.81	3.91	1.25	.71	4.67	.66	.68	4.43	.81	.41
8)	4.07	.96	.73	3.56	1.32	.72	4.67	.55	.59	4.15	.96	.53
9)	4.03	1.09	.68	3.81	1.19	.62	4.70	.54	.72	4.50	.83	.75
10)	4.00	1.10	.61	3.84	1.21	.69	4.37	.85	.78	4.22	1.08	.81
11)	4.21	1.01	.67	3.94	1.11	.69	4.70	.60	.23	4.35	.96	.72
12)	3.86	1.36	.81	3.87	1.16	.76	4.70	.54	.50	4.04	1.06	.71
13)	3.76	1.27	.81	3.74	1.39	.76	4.60	.81	.53	4.17	1.04	.70
14)	4.17	1.00	.49	3.88	1.19	.72	4.43	.90	.44	4.10	.93	.59
15)	4.52	.78	.78	4.08	1.04	.71	4.80	.55	.39	4.30	.96	.55

Appendix B

Examiner Instructions

General Directions

[To follow informal introductory remarks by Examiner] "We appreciate the fact that you're willing to help us by trying out some new kinds of questions. These are in small booklets like this one. [Hold up a sample!]

In most of the booklets there are also pictures along with the questions. Everything should be pretty easy to understand - but I'll read almost all the things written in the booklets to you, so you don't have to worry too much about knowing all the words. A lot of the questions have to do with how you feel about different jobs; what you think people are like; how you feel about yourself and how you feel about the Neighborhood Youth Corps. Let me know if there is anything in them that's putting you down or would be insulting to answer, because you don't have to answer anything that you feel we have no right to ask. But we think the questions are pretty straight and we hope you'll have some fun working out a lot of them.

The answers that you give to any of the questions will never be seen by anybody at an N.Y.C. project and none of this is going to be used to affect anything you do in N.Y.C. So you're free to answer the way you really feel - you don't have to impress anybody."

JOB KNOWLEDGE (JK)

I. ADMINISTRATION

Examiner Reads:

- Instructions
- Job Titles
- Item Stems
- Item choices (Tools pictured are not to be named by the examiner)

II. INSTRUCTIONS

"What we would like to find out with the questions in this booklet is how much you know about different jobs. You'll see pictures of all sorts of jobs and the name of each job - like a 'Dock Worker'(or Longshoreman)."

[Have enrollees turn to first item in booklet.]

"Right after the picture of the Dock Worker are three questions. The first one says: 'How much schooling do you think you'd need to get this job?' If it's a college degree mark the first box. If you need a trade school certificate mark the second box." [Continue reading item stems, item choices and job titles for all items.]

VOCATIONAL PLANS AND ASPIRATIONS (VP, VA)

I. ADMINISTRATION

Examiner Reads:

- Instructions
- Job Titles
- Job Descriptions
- Response Alternatives

II. INSTRUCTIONS

(a) Aspirations:

"In this booklet you'll see pictures of different kinds of jobs that people have. With each picture it tells the name of the job; then it says a little about what kind of work a person does on that job."

"Let's suppose you could have any job you see here, if you wanted it. How would you feel about working at each of these jobs. Like an auto mechanic. You tell how you feel about that job by making a mark in one of the four boxes next to the picture. If you feel its a great job; the kind of thing you would like--draw a line through the first box. If its pretty good and you wouldn't mind doing this--mark the second box. If it's not the greatest, but you would do it if there was nothing better--mark the third box. And if its a rotten kind of job that you wouldn't like at all--mark the last box." [Examiner: Points out each response statement in booklet.]

VOCATIONAL PLANS AND ASPIRATIONS (Continued)

"Now look at this first picture again. It shows a crane operator (Doctor for female booklet) and it says 'Operates a crane to lift heavy things.'"

"Okay, mark the box that shows how you feel about that job."

[Continue with remaining items; reading response alternatives for first several items or until it is apparent that the examinees fully understand the response continuum. Read all job titles and job descriptions.]

(b) Plans:

"Now we're going to look at the same pictures again, but this time I want you to do something a little different. Think about what you might be able to do after you leave the N.Y.C. program and what kind of jobs you would look for over the next few years. Then look at the first picture that shows the crane operator. If you think the odds are that you would do something like that--then mark the first box. If you think there's a fair chance you would wind up in this job--mark the next box. If you think this job is possible, but you don't think chances are too good that you would really do this--then mark the third box. If this is way out for you and you wouldn't even look for this kind of job--put a mark in the last box down."

"Now look at this first picture again. It shows the crane operator and it says, 'Operate a crane to lift heavy things.'"

"Okay, mark the box that shows how you feel about that job."

[Continue with remaining items; reading response alternatives for first several items, or until it is apparent the examinees fully understand these. Read all job titles and job descriptions.]

ATTITUDE TOWARD AUTHORITY (AA)

I. ADMINISTRATION

Examiner Reads

- Instructions
- Verbal material accompanying each pictorial scene
- Item Stems
- Response Alternatives

(Only a simple descriptive statement by the examiner is acceptable in referring to any pictorial content - e.g., "the cop talking to the guy sitting on the grass.")

II. INSTRUCTIONS

"What we would like to find out now is how much you know about the way people would act with each other in different situations. You'll see pictures of people in a lot of everyday scenes - like the first picture in this booklet that shows a guy talking to a teacher who says 'You'll have to do better, your marks aren't so good.' He answers - (Right below the picture) - 'I'll try to study more and get someone to help me with the work.'"

"Now over on the right it says, 'Statistics would show that the number of guys who give this answer to the teacher is: 2 out of 10 guys, 3 out of 10 guys, 4 out of 10, 5 out of 10 or 6 out of 10 guys.'"

"Mark the box that tells best how many guys you think would give this answer to a teacher."

[Read remaining items in the same way.]

SELF-ESTEEM (SE)

I. ADMINISTRATION

Examiner Reads

- Instructions
- Verbal Material in Pictorial Scene
- Response Alternatives

II. INSTRUCTIONS

"What you'll see here are pictures about everyday things that happen to young people. Maybe even some that could happen to you. Like the picture where you see a guy (girl) who's applying for a job and the man at the desk says 'So you think you want a job here!' Make believe that the young person in the picture is you. Below the picture are some of the ways you might feel if it really were you standing there."

"Put a line in the box next to the way that tells best how you would feel. If you would think 'They would probably tell me I'm someone they can't use' - put a line in the first box. If you would think 'They would offer me a pretty good job' - mark the second box (where it says that). If you would think 'They might offer me some kind of job - but nothing too good' mark the third box."

"Now let's look at the next picture."

[Continue reading response alternatives in the same manner as above. The brief description of the pictorial scene is only to be used for purposes of referring to the item, e.g., "The girl standing at the counter" Do not elaborate on the pictorial content unnecessarily.]

INTEREST IN VOCATIONAL TASKS (VI)

I. ADMINISTRATION

Examiner Reads:

- Instructions
- Job Tasks
- Response Alternatives

II. INSTRUCTIONS

"This next set of pictures shows different things people do on jobs. With each picture, it tells what the person is doing. Some of these things you might like to do. Some of them you might not like at all. For some you might not care too much one way or the other."

"For example, look at the first picture where it says 'Be responsible for taking care of office files,' and that's what it shows the guy doing."

"Now think of how you would feel about doing that yourself. Don't worry about whether you know how to do this or if you have enough school or training--just how you feel about doing it. If this is the worst and you couldn't take doing it at all; just mark through the first box (where it says that). If its not just what you would like but not the worst - put a mark through the next box. If it's pretty good and you wouldn't mind doing this - mark the third box down. And if it's great - just what you would like to do, put a mark through the last box."

INTEREST IN VOCATIONAL TASKS (Continued)

"Okay, now look at the next picture, it says. . ."

[Continue as for the first item. Read all tasks: read response alternatives only until examinees readily grasp the four point response scale.]

DEFERRED GRATIFICATION (DG)

I. ADMINISTRATION

Examiner Reads

- Instructions
- Item Stems
- Response Alternatives

II. INSTRUCTIONS

"The things written in this booklet tell some of the ways people feel about getting along in life. We would like to see how good you are at guessing what most people might think about these. I'll read each one with you and you'll tell how you think people would feel about it. Like the first one on the page that says . . . 'It pays to listen to advice your parents give you.'"

"If you think almost nobody would feel this way mark the first box. If you think only a few people might agree with this then mark the second box. If you think some people would agree mark the third box. If a lot of people would feel this way mark the fourth box and if you think most everybody would agree mark the last box. (Mark only one box - the one that's closest to the way you think it is.)"

"Now the next one says . . ."

[Continue for remaining 19 statements. Once the response continuum is apparent to the enrollees it is not necessary to repeat the reading of all 5 response alternatives for every item.]

MOTIVATION FOR VOCATIONAL ACHIEVEMENT (MVA)

I. ADMINISTRATION

Examiner Reads

- Instructions
- Item Stems
- Response Alternatives

II. INSTRUCTIONS

"The things we're asking about here have to do with what you think about jobs and about going to work. So we wrote down what people sometimes say about jobs and about working and we want to know how you feel about these."

"The first one says 'I would go to school at night for special courses if it meant I could get a better job.' Under that it says 'Nearly Always;' 'Most of the time;' 'Sometimes this is so;' 'Very few times;' and 'This is almost never so.'

"Put a line through the box next to the way that tells best how you feel."

[Instruct examinee to respond to the first item. Continue with oral presentation of next 11 items.]

"Now these next ones are a little different. Here in the booklet it says 'Let's say you have a job that you know was going to be a much better one if you stuck it out. How far would you go to stay with it?'"

"Would you stay with it if..."

[Read item stems and responses for each of the remaining seven items.]

JOB SEEKING AND JOB HOLDING SKILLS (JS;JH)

I. ADMINISTRATION

Examiner Reads:

- Instructions
- Job-seeking-skills item stems and responses (Do not read material in want ads or any content of application blanks)
- Job holding skills item stems; item responses and verbal material in pictorial scenes. (Simple description of pictorial scene is permitted - e.g., "boss talking to a guy coming in to work").

II. INSTRUCTIONS

"We're interested in how you feel people should act on a job - and what you might do on a job when certain things happen. For example, it says here [Read first item]. If you think you give them one week's pay - mark the first box. If you give them - whatever you can afford - mark the second box [Read remaining alternatives: Read next verbal item.]

"Now we're going to look at some job 'want ads' - like the kind that you see in the newspaper when someone wants to hire people for jobs. Read each want ad on your own [Allow approximately 30 seconds for reading of ad] - then I'll read the questions to you that ask about what's in the want ad." [Read the nine want ad items.]

"These next questions are about parts of an application blank - like the kind you fill out when you go to get hird for a job. Look at this part of the application blank and the things it asks you to fill out." [Allow approximately 30 to 45 seconds to look over blank. Then read each item stem and alternatives.]

"Now look at the next application blank that some guy (girl) filled in. Read that over." [Allow about one minute to read material. Read each item stem and alternatives.] [Continue on to Job Holding items - read all verbal material including verbal portions of pictorial scenes.]

PRACTICAL REASONING (PR - Z,M,F)

I. ADMINISTRATION

Examiner Reads*

- Instructions as given in each Booklet
- Item Stems
- Item Responses

II. INSTRUCTIONS

"Now we're going to go through questions in three booklets that ask you to do some of the things you might have to do on certain jobs. In one you'll be asked to follow directions for delivering things to people. In another, you'll be asked about sorting mail in a post office using zip codes and in a third, you'll answer questions about filing cards."

"Let's look at this first booklet. Open it up and we'll read the directions."

[Administer Map Reading subtest. Examiner reads directions and items.]

"Now open this next booklet."

[Administer Zip Code subtest.]

"Now open the third booklet. Pull off the envelope on the back of the booklet and take out the cards inside."

[Administer Card Filing subtest.]

---

*Do not read complete details of map, or zip-code chart. The reading of headings only for Practical Reasoning file cards is permissible.

ENROLLEE RATING SCALE (ERS)

I. ADMINISTRATION

Examiner Reads

- Instructions
- Item Stems
- Response Alternatives

II. INSTRUCTIONS

"On the pages of this booklet are some of the things that enrollees say are important about N.Y.C. We would like to know how you think these comments fit the N.Y.C. program. For each comment you tell how much you think it fits your N.Y.C. project by putting a mark in the box next to the way that tells best how you really feel. Like the one on the first page that says . . ."

"If you think this is always the way it is in N.Y.C. mark the first box. If this is pretty much the way it is mark the second box. If it's sometimes so, mark the third box. And if you feel this is never the way it is in your N.Y.C. project, mark the last box.

Appendix C

Counselor and Work Supervisor
Rating Scales

N.Y.C. Project ____________

COUNSELOR'S RATING SCALE

Neighborhood Youth Corps

Counselor's Name ______________________________ Name of Enrollee Being Rated ______________________

On this sheet are 11 statements about enrollee behavior that counselors consider important in determining how well an enrollee is coming along. We would appreciate your evaluation of how these apply to this enrollee.

Please read each statement carefully. Then circle one of the numbers 1 to 5 that best indicates how the statement applies to the enrollee.

1	2	3	4	5
This describes the enrollee perfectly	This is true most of the time	Sometimes this is true of the enrollee	This is not usually so	The enrollee is not like this at all

---

Circle one

1. Pays attention to good grooming and dresses appropriately. 1 2 3 4 5
2. Is not very open about discussing personal and job problems. 1 2 3 4 5
3. Shows a lot of resentment and hostility. 1 2 3 4 5
4. Is cooperative and willing to listen to advice. 1 2 3 4 5
5. Makes realistic plans about future jobs. 1 2 3 4 5
6. Shows little poise or self-assurance. 1 2 3 4 5
7. Is coherent in expressing himself (herself). 1 2 3 4 5
8. Is motivated to want to work and expend effort. 1 2 3 4 5
9. Does not show good day-to-day planning so that he (she) can handle the job (let's home life interfere, for example). 1 2 3 4 5
10. Indicates a willingness to enroll in school or some sort of training on a part-time basis. 1 2 3 4 5
11. Shows qualities that indicate he (she) will do well after leaving N.Y.C. 1 2 3 4 5

N.Y.C. Project ____________

WORK SUPERVISOR'S RATING SCALE

Neighborhood Youth Corps

Work Supervisor's Name ______________________________ Name of Enrollee You are Rating ____________

Below are 10 statements about things that work supervisors consider important when it comes to how the enrollee is doing. We would appreciate your telling us how each one applies to this enrollee.

Please read each statement carefully. Then put a check in one of the five boxes to show how that statement fits the enrollee that you're rating.

1. KNOWS HOW TO FOLLOW INSTRUCTIONS PROPERLY.

This describes just how the enrollee is	This is true most of the time	Sometimes this is true of the enrollee	This is not usually so	The enrollee is not like this at all
☐	☐	☐	☐	☐

2. RESENTS TAKING ORDERS FROM THOSE WHO SUPERVISE HIM (HER).

This describes just how the enrollee is	This is true most of the time	Sometimes this is true of the enrollee	This is not usually so	The enrollee is not like this at all
☐	☐	☐	☐	☐

3. KNOWS HOW TO DRESS RIGHT FOR THE JOB.

This describes just how the enrollee is	This is true most of the time	Sometimes this is true of the enrollee	This is not usually so	The enrollee is not like this at all
☐	☐	☐	☐	☐

4. TAKES SOME PRIDE IN THE WORK AND DOESN'T JUST RUSH THROUGH TO GET IT FINISHED.

This describes just how the enrollee is	This is true most of the time	Sometimes this is true of the enrollee	This is not usually so	The enrollee is not like this at all
☐	☐	☐	☐	☐

5. HAS TO BE TOLD WHAT TO DO EVERY MINUTE OR HE (SHE) CAN'T KEEP BUSY.

This describes just how the enrollee is	This is true most of the time	Sometimes this is true of the enrollee	This is not usually so	The enrollee is not like this at all
☐	☐	☐	☐	☐

6. GETS ALONG WITH OTHERS ON THE JOB.

This describes just how the enrollee is	This is true most of the time	Sometimes this is true of the enrollee	This is not usually so	The enrollee is not like this at all
☐	☐	☐	☐	☐

7. CAN'T GET TO WORK ON TIME.

This describes just how the enrollee is	This is true most of the time	Sometimes this is true of the enrollee	This is not usually so	The enrollee is not like this at all
☐	☐	☐	☐	☐

8. SHOWS SOME INITIATIVE IN TAKING ON A PIECE OF WORK.

This describes just how the enrollee is	This is true most of the time	Sometimes this is true of the enrollee	This is not usually so	The enrollee is not like this at all
☐	☐	☐	☐	☐

9. DOESN'T MAKE TROUBLE ON THE JOB.

This describes just how the enrollee is	This is true most of the time	Sometimes this is true of the enrollee	This is not usually so	The enrollee is not like this at all
☐	☐	☐	☐	☐

10. ASKS QUESTIONS IF PROBLEMS COME UP -- DOESN'T JUST GO AHEAD AND DO THE JOB WRONG.

This describes just how the enrollee is	This is true most of the time	Sometimes this is true of the enrollee	This is not usually so	The enrollee is not like this at all
☐	☐	☐	☐	☐

11. SHOWS THE KIND OF QUALITIES ON THE JOB THAT WILL MAKE HIM A GOOD WORKER AFTER HE LEAVES N.Y.C.

This describes just how the enrollee is	This is true most of the time	Sometimes this is true of the enrollee	This is not usually so	The enrollee is not like this at all
☐	☐	☐	☐	☐